

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OK protein - protein search, using sw model

Run on: September 10, 2003, 00:19:16 ; Search time 86 Seconds

(without alignments)
804.707 Million cell updates/sec

Title: US-09-815-242-12600
Perfect score: 2213
Sequence: 1 MTKPIVAIVGRPMVKSTIF.....IRAFCECTPIHIAKRN 436

Scoring table:
BLOSUM62
Gapop 10.0 , Capept 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

1: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:*
2: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:*
3: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:*
4: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:*
5: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:*
6: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:*
7: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:*
8: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:*
9: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:*
10: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:*
11: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:*
12: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:*
13: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:*
14: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:*
15: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:*
16: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:*
17: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:*
18: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:*
19: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:*
20: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:*
21: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:*
22: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:*
23: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:*
24: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2003.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2213	100.0	436	22	AAU37007
2	2213	100.0	436	22	AAU37139
3	2213	100.0	436	22	AAAB2089
4	2207	99.7	435	22	AAU33787
5	2094	94.6	443	23	ABP8138
6	1731	78.2	436	23	ABP8212
7	1528	69.0	436	23	ABP7827
8	1524	68.9	319	22	AA62833
9	1523	68.8	436	23	ABP27826

10	1522	68.8	436	21	AA804108	YphC protein of St
11	1522	68.8	436	21	AA81770	Streptococcus pneu
12	1522	68.8	436	24	ABU02185	S. pneumoniae type
13	1520	68.7	435	23	ABR54068	Lactococcus lactis
14	993	44.9	448	22	ACG99939	ERA binding domain
15	935.5	42.3	448	22	ACG99941	ERA binding domain
16	923	41.7	503	22	AAU34658	E. coli cellular p
17	907	41.0	501	22	ACG99942	Salmonella typhi c
18	899	40.6	490	22	AAU38110	C. glutamicum prote
19	862.5	39.0	519	22	ACG91323	N. gonorrhoeae aml
20	857.5	38.7	500	24	ABP78407	Haemophilus influe
21	852	38.5	504	22	AAU35376	Haemophilus influe
22	852	38.5	504	22	AA888526	Haemophilus influe
23	852	38.5	504	23	AAU91457	Haemophilus influe
24	851.5	38.5	709	23	ABP65595	Bifidobacterium 10
25	845.5	38.2	461	22	AAU60718	Protonibacterium
26	841.5	38.0	246	22	AAU37925	Streptococcus pneu
27	840.5	37.7	493	22	AAU36362	Pseudomonas aerugi
28	835	37.7	502	22	ACG99943	ERA binding domain
29	799.5	36.1	459	22	ACG99945	ERA binding domain
30	776.5	35.1	459	22	ACG99945	Enterococcus faec
31	754.5	34.1	225	22	AAU33525	Enterococcus faec
32	750.5	33.9	225	22	AAU35096	Amino acid sequenc
33	708.5	32.0	493	20	AAU37461	ERA binding domain
34	707	31.9	456	22	ACG99937	Helicobacter pylor
35	706	31.9	458	22	AAU33777	Amino acid sequenc
36	685.5	31.0	487	20	AAU35489	ERA binding domain
37	669.5	30.3	453	22	ACG99944	ERA binding domain
38	624	28.2	433	22	ACG99948	ERA binding domain
39	620.5	28.0	530	22	ACG99947	ERA binding domain
40	613	27.7	292	22	ACG99940	ERA binding domain
41	559.5	25.3	383	17	AAU77678	ERA binding domain
42	528.5	23.9	331	23	ABU51840	Helicobacter pylor
43	433.5	15.4	106	22	ACG99939	ERA binding domain
44	340	15.4	208	22	ACG99939	ERA binding domain
45	315	14.2	92	19	AAU86046	S. pneumoniae derl

ALIGNMENTS

RESULT 1	
ID	AAU37007
AAU37007	standard; Protein; 436 AA.
AC	AAU37007:
XX	14-FEB-2002 (first entry)
DE	Staphylococcus aureus cellular proliferation protein #1177.
XX	Antisense; prokaryotic cellular proliferation protein;
KW	antibiotic; antibacterial; drug design.
XX	Staphylococcus aureus.
OS	WO200170955-A2.
XX	27-SEP-2001.
PD	21-MAR-2001; 2001WC-US09180.
XX	21-MAR-2000; 2000US-191078P.
PF	23-MAY-2000; 2000US-206848P.
XX	26-MAY-2000; 2000US-207777P.
PR	23-OCT-2000; 2000US-242578P.
PR	27-NOV-2000; 2000US-253625P.
PR	22-DEC-2000; 2000US-257931P.
PR	16-FEB-2001; 2001US-269308P.
XX	(ELIT-) ELITRA PHARM INC.
PA	Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Travick JD, Carr GJ.
XX	
PI	

PI Yamamoto RT, Xu HH;

XX MPI: 2001-611495/70.

DR N-PSDB: AAS54986.

XX New polynucleotides for the identification and development of

PT antibiotics, comprise sequences of antisense nucleic acids -

XX Example 3: Seq ID No 12600: 511pp: English.

XX The invention relates to antisense inhibitors of genes essential to
CC prokaryotic cellular proliferation, their use in identifying the
CC genes, their use in the discovery of novel antibiotics, the essential
CC genes themselves and the encoded proteins. The prokaryotes used are
CC Escherichia coli, Staphylococcus aureus, Salmonella typhi, Klebsiella
CC pneumoniae, Pseudomonas aeruginosa and Enterococcus faecalis. The
CC invention is also useful for the identification of potential new targets
CC for antibiotic development. The antisense nucleic acids can also be used
CC to identify proteins used in proliferation, to express these proteins,
CC and to obtain antibodies capable of binding to the expressed proteins.
CC The proteins can be used to screen compounds in rational drug discovery
CC programmes. The antisense nucleic acid sequence is also useful to screen
CC for homologous nucleic acids which are required for cell proliferation in
CC a wide variety of organisms. The present sequence represents an
CC essential prokaryotic cellular proliferation protein.
CC Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic
CC format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 436 AA:

Query Match 100.0%: Score 2213; DB 22: Length 436;

Best Local Similarity 100.0%: Pred. No. 4,7e-184;

Matches 436; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 MKRPVAVVGRPNVNGKSTIFNRIYGERVSVEDTQVTRDRIRYSSGEMLTJHDFNIIDTGC 60
DB 1 MKRPVAVVGRPNVNGKSTIFNRIYGERVSVEDTQVTRDRIRYSSGEMLTJHDFNIIDTGC 60
OY 61 IEIGDAPFQTOIRAQAIAIDEADVIFVNVNREGLTQSDENVAOILYKSKKPPVLAANK 120
DB 61 IEIGDAPFQTOIRAQAIAIDEADVIFVNVNREGLTQSDENVAOILYKSKKPPVLAANK 120
OY 121 VDNMEMRTDVIDYSLGFGEPYPISSHGSLGDLDDAVVSHFGEEEDDPDEDTIRLSI 180
DB 121 VDNMEMRTDVIDYSLGFGEPYPISSHGSLGDLDDAVVSHFGEEEDDPDEDTIRLSI 180
OY 181 IGRPNVNGKSLVNAIIGEDRVIVSNVAGTTRDAIDTEXSYDQODVLLIDTAGMKRKKGVY 240
DB 181 IGRPNVNGKSLVNAIIGEDRVIVSNVAGTTRDAIDTEXSYDQODVLLIDTAGMKRKKGVY 240
OY 241 ESTEKSYSVLRALKAIERSNVLVYDAEGCIIEDOKRAGVYHNEGKAVVYVNMKWIVE 300
DB 241 ESTEKSYSVLRALKAIERSNVLVYDAEGCIIEDOKRAGVYHNEGKAVVYVNMKWIVE 300
OY 301 KDSKTMMKFEDEVYKFEFOLDAQAIAFVSAKERTRLRLFPYINASSNHRKRVOSSTLN 360
DB 301 KDSKTMMKFEDEVYKFEFOLDAQAIAFVSAKERTRLRLFPYINASSNHRKRVOSSTLN 360
OY 361 EYVTDALSNMPTPTDGRNLNVFYATOVAIERPTVVFVNDVLMHFYKRYLRLENOIRAA 420
DB 361 EYVTDALSNMPTPTDGRNLNVFYATOVAIERPTVVFVNDVLMHFYKRYLRLENOIRAA 420
OY 421 FGEGETPIRIIAAKRN 436
DB 421 FGEGETPIRIIAAKRN 436

RESULT 2
AAU37139
ID AAU37139 standard: Protein: 436 AA.

AC AAU37139;

XX 14-FEB-2002 (first entry)

DE Staphylococcus aureus cellular proliferation protein #1309.

XX Antisense: prokaryotic cellular proliferation protein;

KW antibiotic; antibacterial; drug design.

XX Staphylococcus aureus.

XX WO200170955-A2.

XX 27-SEP-2001.

XX 21-MAR-2001: 2001WO-0509180.

XX 21-MAR-2000: 2000US-191078P.

XX 23-MAY-2000: 2000US-206848P.

XX 26-MAY-2000: 2000US-207272P.

XX 23-OCT-2000: 2000US-242578P.

XX 27-NOV-2000: 2000US-253625P.

XX 22-DEC-2000: 2000US-257931P.

XX 16-FEB-2001: 2001US-269308P.

XX (ELIT-) ELITRA PHARM INC.

XX Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;

XX Yamamoto RT, Xu HH;

XX MPI: 2001-611495/70.

XX N-PSDB: AAS54998.

XX New polynucleotides for the identification and development of

PT antibiotics, comprise sequences of antisense nucleic acids -

XX Example 3: Seq ID No 12732: 511pp: English.

XX The invention relates to antisense inhibitors of genes essential to
CC prokaryotic cellular proliferation, their use in identifying the
CC genes, their use in the discovery of novel antibiotics, the essential
CC genes themselves and the encoded proteins. The prokaryotes used are
CC Escherichia coli, Staphylococcus aureus, Salmonella typhi, Klebsiella
CC pneumoniae, Pseudomonas aeruginosa and Enterococcus faecalis. The
CC invention is also useful for the identification of potential new targets
CC for antibiotic development. The antisense nucleic acids can also be used
CC to identify proteins used in proliferation, to express these proteins,
CC and to obtain antibodies capable of binding to the expressed proteins.
CC The proteins can be used to screen compounds in rational drug discovery
CC programmes. The antisense nucleic acid sequence is also useful to screen
CC for homologous nucleic acids which are required for cell proliferation in
CC a wide variety of organisms. The present sequence represents an
CC essential prokaryotic cellular proliferation protein.
CC Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic
CC format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 436 AA:

Query Match 100.0%: Score 2213; DB 22: Length 436;

Best Local Similarity 100.0%: Pred. No. 4,7e-184;

Matches 436; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 MKRPVAVVGRPNVNGKSTIFNRIYGERVSVEDTQVTRDRIRYSSGEMLTJHDFNIIDTGC 60
DB 1 MKRPVAVVGRPNVNGKSTIFNRIYGERVSVEDTQVTRDRIRYSSGEMLTJHDFNIIDTGC 60
OY 61 IEIGDAPFQTOIRAQAIAIDEADVIFVNVNREGLTQSDENVAOILYKSKKPPVLAANK 120
DB 61 IEIGDAPFQTOIRAQAIAIDEADVIFVNVNREGLTQSDENVAOILYKSKKPPVLAANK 120
OY 121 VDNMEMRTDVIDYSLGFGEPYPISSHGSLGDLDDAVVSHFGEEEDDPDEDTIRLSI 180

```

DB      |||
121 VDNMEMRTVDYDFYSLGFGEPYPSGSHGLGLDLDAVYSHFGEEDEDPYEDTIRLSI 180
OY      181 IGRPNVGSLLVNAIIGEDRVYSNVAGTTRDAIDTEYSYDGDYVLIDTACMRKKGKY 240
DB      181 IGRPNVGSLLVNAIIGEDRVYSNVAGTTRDAIDTEYSYDGDYVLIDTACMRKKGKY 240
OY      241 ESTEKSYSVLRALKAIERSNVVLYIDAEGGIIEDDKRVAGYAHGOKRAVYVYNNKWDYE 300
DB      241 ESTEKSYSVLRALKAIERSNVVLYIDAEGGIIEDDKRVAGYAHGOKRAVYVYNNKWDYE 300
OY      301 KDSKTMKKFEDEVRKEFOFLDYAOIAFVSAKERTRLTLPFYINEASENNKKRVOSSTLN 360
DB      301 KDSKTMKKFEDEVRKEFOFLDYAOIAFVSAKERTRLTLPFYINEASENNKKRVOSSTLN 360
OY      361 EYVTDATSMNPTPTDKGRRLNVFYATQVAIEPTFVYFVNDVLMHFSYKRYLENOIRAA 420
DB      361 EYVTDATSMNPTPTDKGRRLNVFYATQVAIEPTFVYFVNDVLMHFSYKRYLENOIRAA 420
OY      421 FGEGTPIHIIARKRN 436
DB      421 FGEGTPIHIIARKRN 436

RESULT 3
AAB82089 standard; Protein: 436 AA.
XX
AC      AAB82089:
XX
DT      26-JUN-2001 (first entry)
XX
DE      Staphylococcus aureus yphc protein sequence.
XX
KW      yphc; antimicrobial; cytostatic; antiulcer; microbial infection;
KM      gene therapy; vaccine; gastrointestinal carcinoma; gastric ulcer;
KN      gastritis.
XX
OS      Staphylococcus aureus.
XX
PN      MO200123418-A1.
XX
PD      05-APR-2001.
XX
PE      19-SEP-2000; 2000MO-US25566.
XX
PR      28-SEP-1999; 99US-0406968.
XX
PA      (SMIK ) SMITHKLINE BEECHAM CORP.
XX      (SMIK ) SMITHKLINE BEECHAM PLC.
XX
PI      Zalcain M, Blawas S, Burnham MKR, Sylvester D, Mcdevilt D;
PI      Mathle TB;
XX
DR      WPI: 2001-308138/32.
XX      N-PSDB: AAF86461.
XX
PT      Novel yphc polypeptides of Staphylococcus aureus useful for diagnosing
PT      and treating microbial infections, especially infection by
PT      Staphylococcus aureus and Helicobacter pylori.
XX
PS      Claim 1; Page 3; 41pp: English.
XX
CC      The present sequence is the yphc protein of Staphylococcus aureus. The
CC      yphc coding sequence and protein are useful for treating and diagnosing
CC      microbial infections such as infection caused by S. aureus and
CC      Helicobacter pylori. In addition, the yphc coding sequence and protein
CC      are useful for treating diseases such as H. pylori-induced cancers, e.g.
CC      gastrointestinal carcinoma, gastric ulcers, and gastritis. The present
CC      sequence was obtained from a library of clones of chromosomal DNA of
CC      S. aureus in E.coli. The sequencing data from two or more clones
CC      comprising overlapping S. aureus DNAs was used to construct the present
CC      contiguous DNA sequence.

```

```

XX      SQ      Sequence      436 AA;
XX      Query Match      100.0%; Score 2213; DB 22; Length 436;
XX      Best Local Similarity 100.0%; Pred. No. 4,7e-184;
XX      Matches 436; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY      1 MTRPIVIAIGRPVNGKSTIFNRIIGERVSIYEDTPGVTRDRITSSGEMLTNDENIITDGC 60
DB      1 MTRPIVIAIGRPVNGKSTIFNRIIGERVSIYEDTPGVTRDRITSSGEMLTNDENIITDGC 60
OY      61 IETGDAPFQTOIRAQAEIADADYIIFPVNVRREGLTQSDDMAQILYKSKRPVLAIVK 120
DB      61 IETGDAPFQTOIRAQAEIADADYIIFPVNVRREGLTQSDDMAQILYKSKRPVLAIVK 120
OY      121 VDNMEMRTVDYDFYSLGFGEPYPSGSHGLGLDLDAVYSHFGEEDEDPYEDTIRLSI 180
DB      121 VDNMEMRTVDYDFYSLGFGEPYPSGSHGLGLDLDAVYSHFGEEDEDPYEDTIRLSI 180
OY      181 IGRPNVGSLLVNAIIGEDRVYSNVAGTTRDAIDTEYSYDGDYVLIDTACMRKKGKY 240
DB      181 IGRPNVGSLLVNAIIGEDRVYSNVAGTTRDAIDTEYSYDGDYVLIDTACMRKKGKY 240
OY      241 ESTEKSYSVLRALKAIERSNVVLYIDAEGGIIEDDKRVAGYAHGOKRAVYVYNNKWDYE 300
DB      241 ESTEKSYSVLRALKAIERSNVVLYIDAEGGIIEDDKRVAGYAHGOKRAVYVYNNKWDYE 300
OY      301 KDSKTMKKFEDEVRKEFOFLDYAOIAFVSAKERTRLTLPFYINEASENNKKRVOSSTLN 360
DB      301 KDSKTMKKFEDEVRKEFOFLDYAOIAFVSAKERTRLTLPFYINEASENNKKRVOSSTLN 360
OY      361 EYVTDATSMNPTPTDKGRRLNVFYATQVAIEPTFVYFVNDVLMHFSYKRYLENOIRAA 420
DB      361 EYVTDATSMNPTPTDKGRRLNVFYATQVAIEPTFVYFVNDVLMHFSYKRYLENOIRAA 420
OY      421 FGEGTPIHIIARKRN 436
DB      421 FGEGTPIHIIARKRN 436

RESULT 4
AAU33787
ID      AAU33787 standard; Protein: 435 AA.
XX
AC      AAU33787:
XX
DT      14-FEB-2002 (first entry)
XX
DE      Staphylococcus aureus cellular proliferation protein #63.
XX
KW      Antisense; prokaryotic cellular proliferation protein;
KM      antibiotic; antibacterial; drug design.
XX
OS      Staphylococcus aureus.
XX
PN      MO200170955-A2.
XX
PD      27-SEP-2001.
XX
PE      21-MAR-2001; 2001MO-US09180.
XX
PR      21-MAR-2000; 2000US-191078P.
XX      23-MAY-2000; 2000US-206848P.
XX      26-MAY-2000; 2000US-207127P.
XX      23-OCT-2000; 2000US-242578P.
XX      27-NOV-2000; 2000US-253525P.
XX      22-DEC-2000; 2000US-257931P.
XX      16-FEB-2001; 2001US-269508P.
XX
PA      (ELIT-) ELITRA PHARM INC.
XX
PI      Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;
PI      Yamamoto RT, Xu HH;

```

XX WPI: 2001-611495/70.
 DR N-PSDB: AAS51646.
 XX New polynucleotides for the identification and development of
 PT antitoxins, comprise sequences of antisense nucleic acids -
 XX
 PS Example 3: Seq ID No 5283; 511pp; English.
 XX
 CC The invention relates to antisense inhibitors of genes essential to
 CC prokaryotic cellular proliferation, their use in identifying the
 CC genes, their use in the discovery of novel antibiotics, the essential
 CC genes themselves and the encoded proteins. The prokaryotes used are
 CC *Escherichia coli*, *Staphylococcus aureus*, *Salmonella typhi*, *Klebsiella*
 CC *pneumoniae*, *Pseudomonas aeruginosa* and *Enterococcus faecalis*. The
 CC invention is also useful for the identification of potential new targets
 CC for antibiotic development. The antisense nucleic acids can also be used
 CC to identify proteins used in proliferation, to express these proteins,
 CC and to obtain antibodies capable of binding to the expressed proteins.
 CC The proteins can be used to screen compounds in rational drug discovery
 CC programmes. The antisense nucleic acid sequence is also useful to screen
 CC for homologous nucleic acids which are required for cell proliferation in
 CC a wide variety of organisms. The present sequence represents an
 CC essential prokaryotic cellular proliferation protein.
 CC Note: The sequence data for this patent did not form part
 CC of the printed specification, but was obtained in electronic
 CC format directly from Wipo at
 CC ftp.wipo.int/pub/published_pcl_sequences.
 CC
 XX Sequence 435 AA:
 XX
 Query Match 99.7% Score 2207; DB 22; Length 435;
 Best Local Similarity 100.0%; Pred. No. 1.6e-183;
 Matches 435; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MTKPIVAIGRPVNGKSTIFNRIVGERVSIYEDTGVYRDRIVSSGMLTHDNIIDTGC 60
 DB 1 MTKPIVAIGRPVNGKSTIFNRIVGERVSIYEDTGVYRDRIVSSGMLTHDNIIDTGC 60
 QY 61 IEIGDAFFQOIRAOAEIAIDADVIFNVNVRGLOSDEMAOILYKSKKPVYLAUNK 120
 DB 61 IEIGDAFFQOIRAOAEIAIDADVIFNVNVRGLOSDEMAOILYKSKKPVYLAUNK 120
 QY 121 VDNKEMRTDYYDYSGLGCEPPYISGSHGLGDLDAVNSHCEEBEDPYDDETI RLST 180
 DB 121 VDNKEMRTDYYDYSGLGCEPPYISGSHGLGDLDAVNSHCEEBEDPYDDETI RLST 180
 QY 181 IGRPNVGSLLVNAIIGEDRVYVSNVACTRDAIDTEYSYDGDYVLIIDTGRKKKKGYV 240
 DB 181 IGRPNVGSLLVNAIIGEDRVYVSNVACTRDAIDTEYSYDGDYVLIIDTGRKKKKGYV 240
 QY 241 ESTEKYSVLRAKAIERSNVYLVYDAEOGIIEDDKRVAGYAHBOGKAVYVYVNNKDDTVE 300
 DB 241 ESTEKYSVLRAKAIERSNVYLVYDAEOGIIEDDKRVAGYAHBOGKAVYVYVNNKDDTVE 300
 QY 301 KOSKTKMKKFEDEVKREFOPLDYAOIAFVSAKERTRLTLPFYINEASENKKRVOSSTLN 360
 DB 301 KOSKTKMKKFEDEVKREFOPLDYAOIAFVSAKERTRLTLPFYINEASENKKRVOSSTLN 360
 QY 361 EYVTDALISNNPTPTDKGRRLNVFATQVAIEPTFEVFNVDVDELMMHRSYRYLENDIRAA 420
 DB 361 EYVTDALISNNPTPTDKGRRLNVFATQVAIEPTFEVFNVDVDELMMHRSYRYLENDIRAA 420
 QY 421 FCEECTPIHIIARRK 435
 DB 421 FCEECTPIHIIARRK 435

XX 24-JUL-2002 (first entry)
 DT Staphylococcus epidermidis ORF amino acid sequence SEQ ID NO:3183.
 DE Staphylococcus epidermidis: open reading frame; ORF: bacterial infection;
 XX antibacterial; gene therapy.
 XX
 OS Staphylococcus epidermidis.
 XX
 PN US6380370-B1.
 XX
 PD 30-APR-2002.
 XX
 PF 13-AUG-1998; 98US-0134001.
 XX
 PR 14-AUG-1997; 97US-055779P.
 XX
 PR 08-NOV-1997; 97US-064964P.
 XX
 PA (GENO-) GENOME THERAPEUTICS CORP.
 XX
 PI Doucette-Stamm LA, Bush D;
 XX
 DR WPI: 2002-381255/41.
 XX
 DR N-PSDB; ABN90883.
 XX
 PS Disclosure: SEQ ID 3183; 267pp; English.
 XX
 CC ABN90538 to ABN93374 represent *Staphylococcus epidermidis* open reading
 CC frame (ORF) nucleic acid sequences which encode the amino acid sequences
 CC given in ABP5124 to ABP37960. The *S. epidermidis* sequences have
 CC antibacterial activity and can be used in gene therapy. The sequences
 CC can also be used in the diagnosis and treatment of bacterial infections,
 CC particularly *S. epidermidis* infections. The sequences can be used to
 CC screen for compounds able to interfere with the *S. epidermidis* life
 CC cycle or inhibit *S. epidermidis* infection.
 CC N.B. The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from the
 CC USPO web site.
 CC
 XX Sequence 443 AA:
 XX
 Query Match 94.6% Score 2094; DB 22; Length 443;
 Best Local Similarity 93.1%; Pred. No. 1.1e-173;
 Matches 406; Conservative 18; Mismatches 12; Indels 0; Gaps 0;
 QY 1 MTKPIVAIGRPVNGKSTIFNRIVGERVSIYEDTGVYRDRIVSSGMLTHDNIIDTGC 60
 DB 8 MTKPIVAIGRPVNGKSTIFNRIVGERVSIYEDTGVYRDRIVSSGMLTHDNIIDTGC 67
 QY 61 IEIGDAFFQOIRAOAEIAIDADVIFNVNVRGLOSDEMAOILYKSKKPVYLAUNK 120
 DB 61 IEIGDAFFQOIRAOAEIAIDADVIFNVNVRGLOSDEMAOILYKSKKPVYLAUNK 120
 QY 68 IEIGDAFFQOIRAOAEIAIDADVIFNVNVRGLOSDEMAOILYKSKKPVYLAUNK 127
 DB 121 VDNKEMRTDYYDYSGLGCEPPYISGSHGLGDLDAVNSHCEEBEDPYDDETI RLST 180
 QY 128 VDNLEHNDIYDYSGLGCEPPYISGSHGLGDLDAVNSHCEEBEDPYDDETI RLST 187
 DB 128 VDNLEHNDIYDYSGLGCEPPYISGSHGLGDLDAVNSHCEEBEDPYDDETI RLST 187
 QY 181 IGRPNVGSLLVNAIIGEDRVYVSNVACTRDAIDTEYSYDGDYVLIIDTGRKKKKGYV 240
 DB 181 IGRPNVGSLLVNAIIGEDRVYVSNVACTRDAIDTEYSYDGDYVLIIDTGRKKKKGYV 240
 QY 241 ESTEKYSVLRAKAIERSNVYLVYDAEOGIIEDDKRVAGYAHBOGKAVYVYVNNKDDTVE 300
 DB 188 IGRPNVGSLLVNAIIGEDRVYVSNVACTRDAIDTEYSYDGDYVLIIDTGRKKKKGYV 247
 QY 248 ESTEKYSVLRAKAIERSNVYLVYDAEOGIIEDDKRVAGYAHBOGKAVYVYVNNKDDTVE 307
 DB 248 ESTEKYSVLRAKAIERSNVYLVYDAEOGIIEDDKRVAGYAHBOGKAVYVYVNNKDDTVE 307
 QY 301 KOSKTKMKKFEDEVKREFOPLDYAOIAFVSAKERTRLTLPFYINEASENKKRVOSSTLN 360
 DB 308 KOSKTKMKKFEDEVKREFOPLDYAOIAFVSAKERTRLTLPFYINEASENKKRVOSSTLN 367

OY 361 EYVTDALSNMPTDCKRRLNVEFYATVOVAIEPPTFVFNVDLHMFSTRYLENQIRAA 420
 DB 368 EYVTDALSNMPTDCKRRLNVEFYATVOVAIEPPTFVFNVDLHMFSTRYLENQIRAA 427
 OY 421 FGEFGTPIHIIARKR 436
 DB 428 FGEFGTPIHIIARKR 443

RESULT 6

AB849212

ID AB849212 standard; Protein: 436 AA.

AC AB849212;

DT 05-FEB-2002 (first entry)

DE Listeria monocytogenes protein #1916.

KM Antibacterial; gene therapy; vaccine; biosynthesis; biodegradation;
 KM vitamin B12; bacterial infection; disease.

OS Listeria monocytogenes.

XX WO200177335-A2.

XX 18-OCT-2001.

XX 11-APR-2001; 2001WO-FR01118.

XX 11-APR-2000; 2000FR-0004629.

XX (INSP) INST PASTEUR.

PI Buchrieser C, Frangeul L, Couve E, Rusniok C, Fajhl H, Dehoux P;
 PI Dusurget O, Chetoui F, Nedjar H, Glaser P, Kunst F, Cossart P;
 PI Danile J, Goebel W, Kreft J, Kuhn M, Ng E, Vazquez-Boland JA;
 PI Dominguez-Bernal G, Garrido-Garcia P, Tierrez-Martinez A, Amend A;
 PI Chakraborty T, Dommann E, Hain T, Berche P, Chaidit A, Durant L;
 PI Perez-Diaz J, Baquero F, Garcia Del Portillo F, Gomez-Lopez N;
 PI Madueno E, De Pablo B, Wehland J, Kaerst U, Entian K, Hauf J;
 PI Rose M, Voss H;

DR WPI: 2002-010914/01.

PT Genomic sequence for Listeria monocytogenes, useful e.g. for treatment
 PT and prevention of Listeria and related bacterial infections, and
 PT related polypeptides

PS Claim 6: SEQ ID No 1917; 192pp; French.

XX The present invention relates to the genome sequence of Listeria
 CC monocytogenes EGD-e (see ABA03041). The genome sequence and fragments of
 CC it are useful for selecting probes and primers for detecting genes in L.
 CC monocytogenes and related organisms, and for studying genetic
 CC polymorphisms and other genomes. The present sequence is a protein
 CC encoded by the genome sequence of the present invention. Proteins
 CC expressed from the genome sequence are useful for raising specific
 CC antibodies. Identification of L. monocytogenes and related organisms, and
 CC for biosynthesis and biodegradation, especially biosynthesis of Vitamin
 CC B12. The genome sequence and proteins encoded by it are also useful for
 CC selecting compounds that regulate gene expression and cell replication
 CC and modulate L. monocytogenes-related diseases. In addition, the genome
 CC sequence and proteins encoded by it are useful in pharmaceutical and
 CC vaccines compositions for the treatment or prevention of infections by L.
 CC monocytogenes and related organisms.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIFO
 CC at ftp.wifo.int/pub/published_pct_sequences.

XX Sequence 436 AA;

Query Match 78.28; Score 1731; DB 23; Length 436;

Best Local Similarity 73.6%; Pred. No. 4,4e-142;
 Matches 320; Conservative 65; Mismatches 50; Indels 0; Gaps 0;

OY 1 MTKPIVAIYGRPNCKSTFNRIYGERVSIYEDPGYTRRIRYSQSGMLTHDNIITDGG 60
 DB 1 MAKPVVAIYGRPNCKSTFNRIYGERVSIYEDPGYTRRIRYSQSGMLTHDNIITDGG 60
 OY 61 IETGAPPTQIRAOAEIAIDBADYITMVNVRREGLTOSDBMAQILYKSKRPVLAVK 120
 DB 61 IDLSDEPFLQIRAOAEIAIDBADYITITRGVADADQAKILYRSMKPIYLAIK 120
 OY 121 VDNMEKRTVYDFYSLSGFGEPYPIGSHGLGDLDAVYSHGEEDEPDYEDTIRLSI 180
 DB 121 VDNPEMRDQIYDFYSLSGFGEPYPIGSHGLGDLGMDAVRANHFKEEDEEPDPTVYKPSL 180
 OY 181 IGRPNVKSLLNALGDEDRVYVNAVGTDAIDITEXSIOGDDYVLITAGKRRKKGY 240
 DB 181 IGRPNVKSLLNALGDEDRVYVNAVGTDAIDITEXSIOGDDYVLITAGKRRKKGY 240
 OY 241 ESTEKYSVLRAKAIERSNVLYVDAEOGIEEDOKRVAGYAHGOKAVIYVKKDVE 300
 DB 241 ESTEKYSVLRAKAIERSNVLYVDAEOGIEEDOKRVAGYAHGOKAVIYVKKDVE 300
 OY 301 KOSKTKKREDEVREREPFLDYAQIAFSAKEPRTLTPPYINAESEMKKRVOSTLN 360
 DB 301 KOKRTINWTFEDIRDFPLAPLYFVSATKORLNLPRLNOVSDHMSLSKVSMLN 360
 OY 361 EYVTDALSNMPTDCKRRLNVEFYATVOVAIEPPTFVFNVDLHMFSTRYLENQIRAA 420
 DB 361 DIVSDAVANPSPMCKGRKRLKIFYTVOAVKPPFVFNVDLHMFSTRYLENQIRAA 420
 OY 421 FGEFGTPIHIIARKR 435
 DB 421 FGEFGTPIHIIARKR 435

RESULT 7

ABP27827

ID ABP27827 standard; Protein: 436 AA.

XX ABP27827;

XX 02-JUL-2002 (first entry)

DE Streptococcus polypeptide SEQ ID NO 4830.

KM Streptococcus: GAS; GBS; group B streptococcus; Streptococcus agalactiae;

KM group A streptococcus; Streptococcus pyogenes; antibacterial;
 KM antinflammatory; infection; vaccine; meningitis; gene therapy.

OS Streptococcus pyogenes.

XX WO200234771-A2.

XX 02-NOV-2002.

XX 29-OCT-2001; 2001WO-GB04789.

XX 27-OCT-2000; 2000GB-0026333.

XX 24-NOV-2000; 2000GB-0028727.

XX 07-MAR-2001; 2001GB-0005640.

PA (CHIR-) CHIRON SPA.

PI Telford J, Maignani V, Margalit Ros YI, Grandi G, Fraser C;

XX WPI: 2002-352536/38.

XX N-PDB: ABN8458.

PT New Streptococcus protein for the treatment or prevention of infection
 PT or disease caused by Streptococcus bacteria, such as meningitis, and

PT for detecting a compound that binds to the protein -
 XX
 XX
 XX
 PS Claim 1: Page 3645: 4525pp: English.

CC The invention relates to a protein (ABP25413-ABP30895) from group B
 CC Streptococcus/GBS (Streptococcus agalactiae) or group A streptococcus/GAS
 CC (Streptococcus pyogenes), comprising one of 5483 sequences (S1), given in
 CC the specification. The proteins have antibacterial and anti-inflammatory
 CC activity. (II), nucleic acids encoding (I), ABP66044-ABP71326 and
 CC antibodies that bind (I) are used in the manufacture of medicaments for
 CC the treatment or prevention of infection or disease caused by
 CC Streptococcus bacteria, particularly S. agalactiae and S. pyogenes.
 CC Nucleic acids encoding (I) are used to detect Streptococcus in a
 CC biological sample. (I) is used to determine whether a compound binds to
 CC (I). A composition comprising (I) or a nucleic acid encoding (I), may be
 CC used as a vaccine or diagnostic composition. The disease caused by
 CC Streptococcus that is prevented or treated may be meningitis. Nucleic
 CC acid encoding (I) may be used to recombinantly produce (I) and may be
 CC used in gene therapy. Antibodies to (I) are used for affinity
 CC chromatography, immunoassays, and distinguishing/identifying
 CC Streptococcus proteins.

XX Sequence 436 AA:

Query Match 69.0%; Score 1528; DB 23; Length 436;
 Best Local Similarity 66.1%; Pred. No. 2,1e-124;
 Matches 288; Conservative 75; Mismatches 71; Indels 2; Gaps 2;

QY 1 MTRPIAIVGRPNVSKSTIFNRIVGERVSIEDTGVYTRDRITSSGEMLTNIDITG 60
 DB 1 MVLPIAIVGRPNVSKSTIFNRIVGERVSIEDTGVYTRDRITSSGEMLTNIDITG 60
 QY 61 IEGAPPTOTIRBAQEAIAIDADVIFMVNVRGLNOSQENVAQILYKSKRVYLA 120
 DB 61 IDVDVAFMEQDKHQAQIAEDADVIFVVGSGCVDADEYKSKIRNTVYLA 120
 QY 121 VDNEMKRTDVFYSLGFGPEYPISSGSLGLDGLDAVYSHRGEEDDPYDEDTIRLSI 180
 DB 121 VDNEMKRTDVFYSLGFGPEYPISSGSLGLDGLDAVYSHRGEEDDPYDEDTIRLSI 180
 QY 122 VDNEMKRTDVFYSLGFGPEYPISSGSLGLDGLDAVYSHRGEEDDPYDEDTIRLSI 179
 DB 122 VDNEMKRTDVFYSLGFGPEYPISSGSLGLDGLDAVYSHRGEEDDPYDEDTIRLSI 179
 QY 181 ICRPNVSKSLVNAIIGEDRVYVNVAGTTTDAIDTETS-VYDQDYLITDTAGRRKGV 239
 DB 181 ICRPNVSKSLVNAIIGEDRVYVNVAGTTTDAIDTETS-VYDQDYLITDTAGRRKGV 239
 QY 180 ICRPNVSKSLVNAIIGEDRVYVNVAGTTTDAIDTETS-VYDQDYLITDTAGRRKGV 239
 DB 180 ICRPNVSKSLVNAIIGEDRVYVNVAGTTTDAIDTETS-VYDQDYLITDTAGRRKGV 239
 QY 240 YESTEKYSVLRAKAIERSNVLVYDAEGCIIDQDRVAGAHQCAAVYVNNKMDV 299
 DB 240 YESTEKYSVLRAKAIERSNVLVYDAEGCIIDQDRVAGAHQCAAVYVNNKMDV 299
 QY 300 EKDSKTKMKFEDEVRKEFOFLDYAQAIAFVSAKERTRLRTLPYINASENNKRRVQS 359
 DB 300 EKDSKTKMKFEDEVRKEFOFLDYAQAIAFVSAKERTRLRTLPYINASENNKRRVQS 359
 QY 360 NEVYTDATISNPTPTDGRRLNFYATQVAIEPTFFVFNVDVLMHFSYKRYLENOIA 419
 DB 360 NEVYTDATISNPTPTDGRRLNFYATQVAIEPTFFVFNVDVLMHFSYKRYLENOIA 419
 QY 420 AFGFEGCTPIHITARRK 435
 DB 420 AFGFEGCTPIHITARRK 435
 QY 435 AFGFEGCTPIHITARRK 435
 DB 435 AFGFEGCTPIHITARRK 435

RESULT 8
 AACG2833
 ID AACG2833 standard; Protein: 319 AA.

XX AACG2833:
 XX
 XX 03-SEP-2001 (first entry)
 XX
 XX S. epidermidis open reading frame protein sequence SEQ ID NO:2760.
 XX
 XX Staphylococcus epidermidis SRI strain; Infection: diagnosis;
 KM vaccination: endocarditis.

XX
 OS Staphylococcus epidermidis.
 XX
 XX
 PN W0200134809-A2.
 XX
 XX 17-MAY-2001.
 PD
 XX
 XX 09-NOV-2000; 2000MO-US30782.
 PF
 XX
 XX 09-NOV-1999; 99US-0164258.
 PR
 XX
 PA (GLAX) GLAXO GROUP LTD.
 XX
 XX
 P1 Kimerly MJ;
 XX
 DR WPI: 2001-316495/33.
 XX
 DR N-PSDB: AAH53683.
 XX
 XX
 PS Claim 18: Page 719: 2188pp: English.

CC AAH52304 to AAH53970 represent nucleic acids (I) encoding polypeptides
 CC (I), given in AAH5154 to AAH53120, from Staphylococcus epidermidis.
 CC (I) and (II) can have antibacterial activity and therefore can be used
 CC in vaccination. The nucleic acids (I) may be used to produce the
 CC S. epidermidis polypeptides (II) via the production of vectors
 CC containing them which are used to produce hosts cells which express the
 CC polypeptides. The polypeptides (II) (and/or nucleic acids) may then be
 CC used to vaccinate subjects and to raise antibodies against the bacteria.
 CC The polypeptides may also be used to assay for other inhibitors of their
 CC activity and therefore identify compounds that may be used for the
 CC treatment of S. epidermidis infections, e.g. endocarditis. AAH53971 to
 CC AAH55090 represent specifically claimed S. epidermidis genomic DNA
 CC polynucleotide sequences from the present invention. AAH55091 to
 CC AAH55098 represent oligonucleotide sequences and primers which are used
 CC in the amplification of the present invention.

CC N.B. The present invention specifically claims all the polynucleotide
 CC sequences given in the sequence listing of the present specification,
 CC however the sequence listing only goes up to SEQ ID NO:4454 so even
 CC though sequences are given in the disclosure for SEQ ID NO:4455 to 4472,
 CC no sequences are present for SEQ ID NO:4455 to 4464.

XX Sequence 319 AA:

Query Match 68.9%; Score 1524; DB 22; Length 319;
 Best Local Similarity 92.5%; Pred. No. 3e-124;
 Matches 295; Conservative 12; Mismatches 12; Indels 0; Gaps 0;

QY 118 VNRVNNKEMRTDVFYSLGFGPEYPISSGSLGLDGLDAVYSHRGEEDDPYDEDTIR 177
 DB 118 VNRVNNKEMRTDVFYSLGFGPEYPISSGSLGLDGLDAVYSHRGEEDDPYDEDTIR 177
 QY 1 VNRVNNKEMRTDVFYSLGFGPEYPISSGSLGLDGLDAVYSHRGEEDDPYDEDTIR 60
 DB 1 VNRVNNKEMRTDVFYSLGFGPEYPISSGSLGLDGLDAVYSHRGEEDDPYDEDTIR 60
 QY 178 LSTIIGRPNVKSSLVNAIIGEDRVYVNVAGTTTDAIDTETS-VYDQDYLITDTAGRRK 237
 DB 178 LSTIIGRPNVKSSLVNAIIGEDRVYVNVAGTTTDAIDTETS-VYDQDYLITDTAGRRK 237
 QY 61 LSTIIGRPNVKSSLVNAIIGEDRVYVNVAGTTTDAIDTETS-VYDQDYLITDTAGRRK 120
 DB 61 LSTIIGRPNVKSSLVNAIIGEDRVYVNVAGTTTDAIDTETS-VYDQDYLITDTAGRRK 120
 QY 238 KYESTEKYSVLRAKAIERSNVLVYDAEGCIIDQDRVAGAHQCAAVYVNNKMD 297
 DB 238 KYESTEKYSVLRAKAIERSNVLVYDAEGCIIDQDRVAGAHQCAAVYVNNKMD 297
 QY 121 KYESTEKYSVLRAKAIERSNVLVYDAEGCIIDQDRVAGAHQCAAVYVNNKMD 180
 DB 121 KYESTEKYSVLRAKAIERSNVLVYDAEGCIIDQDRVAGAHQCAAVYVNNKMD 180
 QY 298 TYEKDSKTKMKFEDEVRKEFOFLDYAQAIAFVSAKERTRLRTLPYINASENNKRRVQS 357
 DB 298 TYEKDSKTKMKFEDEVRKEFOFLDYAQAIAFVSAKERTRLRTLPYINASENNKRRVQS 357
 QY 181 TYEKDSKTKMKFEDEVRKEFOFLDYAQAIAFVSAKERTRLRTLPYINASENNKRRVQS 240
 DB 181 TYEKDSKTKMKFEDEVRKEFOFLDYAQAIAFVSAKERTRLRTLPYINASENNKRRVQS 240
 QY 358 TLEEVYTDATISNPTPTDGRRLNFYATQVAIEPTFFVFNVDVLMHFSYKRYLENOI 417
 DB 358 TLEEVYTDATISNPTPTDGRRLNFYATQVAIEPTFFVFNVDVLMHFSYKRYLENOI 417
 QY 241 TLEEVYTDATISNPTPTDGRRLNFYATQVAIEPTFFVFNVDVLMHFSYKRYLENOI 300
 DB 241 TLEEVYTDATISNPTPTDGRRLNFYATQVAIEPTFFVFNVDVLMHFSYKRYLENOI 300
 QY 418 RAAFEGCTPIHITARRK 436
 DB 418 RAAFEGCTPIHITARRK 436
 QY 301 RAAFEGCTPIHITARRK 319
 DB 301 RAAFEGCTPIHITARRK 319

CC used to generate antibodies.

XX Sequence 436 AA:

Query Match 68.8%; Score 1522; DB 21; Length 436;
Best Local Similarity 66.2%; Pred. No. 7.1e-124;
Matches 290; Conservative 73; Mismatches 69; Indels 6; Gaps 3;

QY 1 MTRPIAIVGRPNVGNKSTIFNRIVGRVSIVEDPGVTRDRISGEMLTIDHPIIDTGG 60
DB 1 MALPTIAIVGRPNVGNKSTIFNRIVGRVSIVEDPGVTRDRISGEMLTIDHPIIDTGG 60
QY 61 IEIGDAPFOVOIRAOEIAIDEDVYIFVWVNRREGTOSDEMVAOILYKSKRPVLAVNK 120
DB 61 IDVDAPFMEQIKHQEIAEADVYIFVWVNRREGTOSDEMVAOILYKSKRPVLAVNK 120
QY 121 VDNMEKRTDVFYSLGCEPPYISGSHGLGDLDAVSHFEE--EEDPYDEDTIRL 178
DB 121 VDNPEKRNIDYDFYALGLCEPLISSVHGIGTGVDLAIVENLNEVEEENP--DVYKF 177
QY 179 SIIGRPNVGKSSLVNAILGEDRVIVSNVAGTTRDAIDTEYS-YGGOVVLIDTAGMKRG 237
DB 178 SLIGRPNVGKSSLVNAILGEDRVIVSNVAGTTRDAIDTEYS-YGGOVVLIDTAGMKRG 237
QY 238 KYESTEKYSVLRAKAIERSNVVLYVIDAEGITIEDOKRVAGAHGOKRAVYIVVKNKD 297
DB 238 KYENTEKYSVMRAMAIDRSVDVLYVNAEGIREYDKRIAGFAHGEKGMITIVVKNKD 297
QY 298 TVERKSKTKMKFEDEVKKEFOFLDYAOIAFVSAKERTRLTFPYINEASENHKKRVQSS 357
DB 298 TLEKDNHTKMKNEEDIREQOYLPAPIIFVSALTOKRLHKLPEMIKOISESOTRIPSA 357
QY 358 TLANEVYDAISNMPPTDOKGRNLNFYATOVAIEPPFVNVNDVLEHFSYKRYLENOI 417
DB 358 VLVNDVMDALAINPPTDOKGRKLIFVATOVATKPTPIVIVNEELMHFSYKRYLENOI 417
QY 418 RAAFGEGTPIHIIAKRR 435
DB 418 RKAFFEGTPIHIIAKRR 435

RESULT 11
AAV81730
ID AAV81730 standard; Protein: 436 AA.

AAV81730:
02-JUN-2000 (first entry)
Streptococcus pneumoniae protein sequence ID36.
XX Streptococcus pneumoniae infection; immunogen; antigen; diagnosis; AIDS;
KW bacterial pneumonia; asplenia; heart disease; lung disease; alcoholism;
KW kidney disease; diabetes; immunosuppressive disorder; otitis media;
KW pneumococcal septicemia; sinusitis; meningitis; therapy.
OS Streptococcus pneumoniae.

XX WO200006738-A2.
XX 10-FEB-2000.
XX 27-JUL-1999; 99WO-GB02452.
XX 27-JUL-1998; 98GB-0016336.
XX 19-MAR-1999; 99US-0125329.
XX (MICK-) MICROBIAL TECHNIQS LTD.
XX la Page RMF, Wells JM, Hanniffy SB, Hansbro PM;
XX MPI; 2000-195301/17.
XX N-PSDB; AA291826.
XX DR

XX Streptococcal proteins and polynucleotides useful for diagnosis,
PT treatment and prophylaxis of bacterial infections
XX Claim 2; Page 49; 76pp; English.

This sequence represents a Streptococcus pneumoniae protein of the
CC invention. The proteins (or their homologues, derivatives and/or
CC fragments) are useful as immunogens or antigens. Immunogenic or antigenic
CC compositions comprising the proteins are useful as vaccines and also in
CC diagnostic assays. The sequences are useful for the detection or
CC diagnosis of S. pneumoniae infection, by contacting a sample to be tested
CC with them. Agents capable of antagonising, inhibiting or interfering with
CC the function or expression of the protein or polypeptide are useful in
CC medical compositions in the treatment or prophylaxis of S. pneumoniae
CC infection. As the sequences can be used to treat S. pneumoniae infection,
CC they can be used to treat bacterial pneumonia, which has high rates in
CC young children, the elderly, and in patients with predisposing conditions
CC such as asplenia, heart, lung and kidney disease, diabetes, alcoholism,
CC or with immunosuppressive disorders, especially AIDS. They can also be
CC used to treat pneumococcal septicemia, otitis media, sinusitis, and
CC meningitis.

Sequence 436 AA:

Query Match 68.8%; Score 1522; DB 21; Length 436;
Best Local Similarity 66.2%; Pred. No. 7.1e-124;
Matches 290; Conservative 73; Mismatches 69; Indels 6; Gaps 3;

QY 1 MTRPIAIVGRPNVGNKSTIFNRIVGRVSIVEDPGVTRDRISGEMLTIDHPIIDTGG 60
DB 1 MALPTIAIVGRPNVGNKSTIFNRIVGRVSIVEDPGVTRDRISGEMLTIDHPIIDTGG 60
QY 61 IEIGDAPFOVOIRAOEIAIDEDVYIFVWVNRREGTOSDEMVAOILYKSKRPVLAVNK 120
DB 61 IDVDAPFMEQIKHQEIAEADVYIFVWVNRREGTOSDEMVAOILYKSKRPVLAVNK 120
QY 121 VDNMEKRTDVFYSLGCEPPYISGSHGLGDLDAVSHFEE--EEDPYDEDTIRL 178
DB 121 VDNPEKRNIDYDFYALGLCEPLISSVHGIGTGVDLAIVENLNEVEEENP--DVYKF 177
QY 179 SIIGRPNVGKSSLVNAILGEDRVIVSNVAGTTRDAIDTEYS-YGGOVVLIDTAGMKRG 237
DB 178 SLIGRPNVGKSSLVNAILGEDRVIVSNVAGTTRDAIDTEYS-YGGOVVLIDTAGMKRG 237
QY 238 KYESTEKYSVLRAKAIERSNVVLYVIDAEGITIEDOKRVAGAHGOKRAVYIVVKNKD 297
DB 238 KYENTEKYSVMRAMAIDRSVDVLYVNAEGIREYDKRIAGFAHGEKGMITIVVKNKD 297
QY 298 TVERKSKTKMKFEDEVKKEFOFLDYAOIAFVSAKERTRLTFPYINEASENHKKRVQSS 357
DB 298 TLEKDNHTKMKNEEDIREQOYLPAPIIFVSALTOKRLHKLPEMIKOISESOTRIPSA 357
QY 358 TLANEVYDAISNMPPTDOKGRNLNFYATOVAIEPPFVNVNDVLEHFSYKRYLENOI 417
DB 358 VLVNDVMDALAINPPTDOKGRKLIFVATOVATKPTPIVIVNEELMHFSYKRYLENOI 417
QY 418 RAAFGEGTPIHIIAKRR 435
DB 418 RKAFFEGTPIHIIAKRR 435

RESULT 12
ABU02185
ID ABU02185 standard; Protein: 436 AA.

ABU02185:
11-FEB-2003 (first entry)
S. pneumoniae type 4 strain protein from coding region #1762.
KW Bacterial meningitis; pneumonia; sepsis; otitis media;

ear infection; antiinflammatory; antibacterial; immunostimulant;
 auditory; respiratory; gene therapy; vaccine.
 Streptococcus pneumoniae type 4 strain.
 WO200277021-A2.
 03-OCT-2002.
 27-MAR-2002: 2002MO-IB02163.
 27-MAR-2001: 2001GB-0007658.
 (CHIR-) CHIRON SPA.
 (GENO-) INST GENOMIC RES.
 Masignani V, Tettelin H, Fraser C;
 NPI: 2003-040579/03.
 N-PSDB: ABX07474.
 New proteins and nucleic acid molecules from Streptococcus pneumoniae,
 useful as medicaments for treating or preventing a disease or infection
 due to streptococcus bacteria, such as pneumonia, sepsis, otitis media
 or ear infection
 Claim 1: SEQ ID NO 3534: 56pp; English.
 The invention relates to a protein comprising or having at least 508
 amino acid residues, identified in the
 specification (available on a computer readable format), or its fragment,
 expressed from 2469 identified DNA coding regions from the
 Streptococcus pneumoniae type 4 strain genomic sequence appearing as
 AB556454. Also included are an antibody which binds one of the
 proteins, treating a patient by administering the protein, DNA or
 antibody (in a composition), a kit comprising first and second primers,
 which are the nucleic acid cited above or fragments between nucleotides
 8-100 of a sequence not defined in the specification, for amplifying a
 target sequence contained within a Streptococcus nucleic acid sequence,
 where the first primer is substantially complementary to the target
 sequence and the second primer is substantially complementary to the
 complement of the target sequence, and where the parts of the primers
 having substantial complementarity define the termini of the target
 sequence to be amplified, assay comprising contacting a test compound
 with the protein, and determining whether the test compound binds to the
 protein and a Streptococcus pneumoniae bacterium, where one or more
 genes encoding the proteins has been rendered inactive. The proteins,
 nucleic acid molecules, antibody and compositions are useful as
 medicaments for treating or preventing a disease or infection due to
 streptococcus bacteria, particularly S. pneumoniae, such as pneumonia,
 sepsis, otitis media or ear infection. They are also useful in developing
 vaccines, diagnostics and antibiotics. The methods are useful for
 identifying immunodominant proteins. The present sequence is one of
 the 2469 proteins expressed by the identified coding regions from the
 genomic sequence.
 Note: The sequence data for this patent did not form part
 of the printed specification, but was obtained in electronic
 format directly from Wipo at
 ftp.wipo.int/pub/published_pct_sequences.
 Sequence 436 AA:
 Query Match 68.8%; Score 1522; DB 24; Length 436;
 Best Local Similarity 66.2%; Pred. No. 7.1e-124;
 Matches 290; Conservative 73; Mismatches 69; Indels 6; Gaps 3;
 OY 1 MTRPIAIVGRPNVSKSTFIRIVGERVSYEDTGVTRDRITSSGELTHDFNITDGG 60
 DB 1 MALPTAIAGRPVGVSTLFRNIRAGERSIYEDVAGYVRDRITVATGELNRSFMDITGG 60
 OY 61 IELGDAFPGOTIRAOEAIDADVIYIRNVNREGLTSDDEMAOILYKSKRPVLAANK 120
 DB 61 IDVDAPFMDQIKHQAIMEADAVIVFVSGKEGIDTADDEYVARKLKTHKPVILAVNK 120

OY 121 VDNEMHRTVDYDFSLGCEPPEYISGSHGLGLDLDAVSHFCE--EEDPYDEDTIRL 178
 DB 121 VDNEMHRTVDYDFSLGCEPPEYISGSHGLGLDLDAVSHFCE--EEDPYDEDTIRL 178
 OY 179 SIIRPNVGSLSLVNMGEDRVIVSNVAGTTRDAIDREYS-YGGQDVLVLTGAKKKG 237
 DB 178 SLIRPNVGSLSLVNMGEDRVIVSNVAGTTRDAIDREYS-YGGQDVLVLTGAKKKG 237
 OY 238 KYESTERYSLRALKAERSNVYLVDAEGGIIEDOKRVAGVAHEGKAANVIVKMD 297
 DB 238 KYESTERYSLRALKAERSNVYLVDAEGGIIEDOKRVAGVAHEGKAANVIVKMD 297
 OY 298 TVEKDSKTKMKEDEVEKREFQFLDVAQIAFVSAKERTRLTLPYINASESNHKKRQSS 357
 DB 298 TLEKDNHTKMKMEEDIEQFQYLPVAPPIIFGALTKQRLHKLPEMKINISSEQTRIPSA 357
 OY 358 TLNEVYTDALSMNPTPDGKRLNVEYATQVAIEPTTYVYNDVYELMHFSYKRLENOI 417
 DB 358 VLNDVINDALAINPTPDGKRLNVEYATQVAIEPTTYVYNDVYELMHFSYKRLENOI 417
 OY 418 RAARFECTPIHITARRK 435
 DB 418 RKAARFECTPIHITARRK 435
 RESULT 13
 ID AB854068 standard; Protein: 436 AA.
 AC AB854068;
 DT 16-MAY-2002 (first entry)
 DE Lactococcus lactis protein yphL.
 KW Biosynthesis; biodegradation; lactic bacterium; yogurt; cheese.
 OS Lactococcus lactis IL1403.
 XX FR2807446-A1.
 XX 12-OCT-2001.
 PD 11-APR-2000; 2000FR-0004630.
 PF 11-APR-2000; 2000FR-0004630.
 PR 11-APR-2000; 2000FR-0004630.
 PA (INRG) INRA INST NAT RECH AGRONOMIQUE.
 PI Biotect A, Sorokline A, Renault P, Ehrlich SD;
 DR WPI: 2002-043418/06.
 PT New nucleotide sequence useful in the identification or Lactococcus
 PT Lactis and related species -
 PS Claim 6: SEQ ID NO 770: 2504pp; French.
 The present invention is related to a Lactococcus lactis nucleotide
 sequence (AB854068) and related proteins (AB853300-AB855621). The
 nucleic acid sequence is useful in the detection and/or amplification of
 related species. The proteins of the invention are useful for the
 biosynthesis or biodegradation of a composition of interest. The
 invention helps research in lactic bacteria, particularly useful in the
 production of yogurt and cheese.
 Note: The sequence data for this patent is based on equivalent patent
 WO200177334 (published 18-OCT-2001) which is available in electronic
 format directly from Wipo at ftp.wipo.int/pub/published_pct_sequences.
 Sequence 436 AA:

Query Match	68.7%	Score 1520	DB 23	Length 436
Best Local Similarity	66.3%	Pred. No. 1.1e-123		
Matches 289	Conservative 69	Mismatches 76	Indels 2	Gaps 2

[illegible]

XX	RESULT 14
XX	AAG99939
ID	AAG99939 standard; Protein; 448 AA.
XX	
AC	AAG99939:
DT	27-SEP-2001 (first entry)
DE	
XX	ERA binding domain polypeptide SEQ ID NO 381.
XX	
KM	ERA binding domain; <i>Escherichia coli</i> ; GTPase; antimicrobial;
KM	antibacterial; antibiotic; pathogenesis; infection; vaccine;
KM	peptide therapy.
XX	
OS	<i>Mycobacterium leprae</i> .
XX	
PN	WO200153458-A2.
XX	
PD	26-JUL-2001.
XX	
PF	17-JAN-2001; 2001WO-US01786.
XX	
PR	18-JAN-2000; 2000US-0176870.
XX	
PA	(SMIK) SMITHKLINE BEECHAM CORP.
PA	(SMIK) SMITHKLINE BEECHAM PLC.
XX	
PI	Lupas AN, Pearce KH;
XX	
DR	WPI: 2001-476108/51.
XX	
PT	New ERA binding domain polypeptides and polynucleotides encoding them
PT	useful as research reagents and materials for discovery of treatments
PT	and diagnostics for diseases, or for genetic immunisation -
XX	
OS	Claim 1: Page 52-53; 279pp; English.

XX The present invention relates to ERA binding domain polypeptides
CC
CC (AGG9555-AMG99989 and AAM00010-AAM00021). The era gene in *Escherichia*
CC coli codes for an essential GTPase protein able to autophosphorylate at
CC serine and/or threonine residues. The protein has potential antimicrobial
CC and antibacterial activity and is useful in screening for antagonists,
CC agonists and for compounds with antibiotic activity. The proteins are
CC also useful in determining their role in pathogenesis of infection,
CC dysfunction and disease and could be used as part of a vaccine and/or
CC peptide therapy.
XX
XX Sequence 448 AA:

Query Match	199: Conservative	93:	Mismatches	136:	Indels	8:	Gaps	6
Query Match Similarity	44.9%:	Score	993:	DB	22:	length	448:	
Best Local Similarity	45.6%:	Pred.	NO. 8.7e-78:					
Matches	199:	Conservative	93:	Mismatches	136:	Indels	8:	Gaps
Oy	1	MKKPRLVYVGRPNVSGTIFNRIVGERSVIYEDRGVTRDRIYSSGEMVLIHDFNITITGG	60					
	11			11	11		11	
Db	1	MSLPVYALIGCPNPGSTFVYNRLAGNQAALYHDDPGITRRKY - PAFNRHDFPVVDITGG	58					
Oy	61	IEIG - DAFQTOYIQAQAEIADIDADYIIFVNVNREGLTOSDEKVAOILYKSKKPVYLVN	119					
	11	11	11	11	11	11	11	11
Db	59	LVFNDSSEFLPEIRIEQANLA - EAKAALFFVVDGCGCFPASADEIKOMLRDOOSVYILAVN	116					
Oy	120	KVDNMEN - RNDVVDVDFSLGCFGPYPISGSHGLGIGLDLDAVVSHEGSEEDPDVEDYIRL	178					
	11	11	11	11	11	11	11	11
Db	117	KESPDGQAIQAALAEFHHILGLGSEPVPMASHGSGDGLDLALLEYLPAPQEEP - DEDEIKV	175					
Oy	179	STIGCPNPGSSLYVNAIIGEDRVIYVSNAGTTRALIDTESVSDGDDVYLITDAPMRKKK	238					
	11	11	11	11	11	11	11	11
Db	176	ALVGRPNVNGSSLLMLTGEQALYISPSGTTTRALIDMYERNKGKYLITDIAIRKKRN	235					
Oy	239	YVESTERTSVLRALKALERSNVVLYIDAEGGITEDDCKRVAGYAHGCKAVYLYVNNKDT	298					
	11	11	11	11	11	11	11	11
Db	236	VDYGAEEFGINRAFEKAIIRADVAVLCVLDGCVTDEDLKLAGIILEDRAVVALYNKMDA	295					
Oy	299	VEKDSKTMKRFEEDEVRKEFOPLDYAOIAFYSAKERTRLTPRYINEASENNHKRVOST	358					
	11	11	11	11	11	11	11	11
Db	296	VEKDSYTYIEHREQLMRLRYLFMDMAEMLFVSAQGTGLRVOKILDCVDIAAOEHRRRRYTAV	355					
Oy	359	LINEVYTDASINPPIDK - GRRLNVEFYVQALIEPPFVFNVDVLEHMFSSKRYLENOI	417					
	11	11	11	11	11	11	11	11
Db	356	INEVLEEVASHMSRPRTTGKCGKLYGTQVSTQRPALALFVNDNRRNDMYRYTIEKOF	415					
Oy	418	RAAFGEFGTRIIHAR	433					
	11	11	11	11	11	11	11	11
Db	416	RKQGLDFGFSPTRLMR	431					

RESULT	15
AC	AMG99941
AC	AMG99941 standard; Protein: 448 AA.
AC	AMG99941;
DT	27-SEP-2001 (first entry)
XX	
DE	ERA binding domain polypeptide SEQ ID NO 383.
XX	
KW	ERA binding domain; <i>Escherichia coli</i> ; GTPase; antimicrobial;
KW	antibacterial; antibiotic; pathogenesis; infection; vaccine;
KW	peptide therapy.
XX	
OS	<i>Treponema pallidum</i> .
XX	
PN	MO200153458-A2.
XX	
PD	26-JUL-2001.
XX	
PF	17-JAN-2001; 2001WO-US01786.
XX	
PR	18-JAN-2000; 2000US-0176870.
XX	

xx (SMK) SMITHKLINE BEECHAM CORP.
 PA (SMK) SMITHKLINE BEECHAM PLC.
 xx
 PI Lupus AN, Pearce KH;
 xx
 DR WPI: 2001-476108/51.
 xx

PT New ERA binding domain polypeptides and polynucleotides encoding them,
 PT useful as research reagents and materials for discovery of treatments
 PT and diagnostics for diseases, or for genetic immunisation .
 xx

Claim 1; Page 53; 279pp; English.

xx The present invention relates to ERA binding domain polypeptides
 CC (AAG9559-AAG9989 and AAM00010-AAM00021). The era gene in Escherichia
 CC coli codes for an essential GTPase protein able to autophosphorylate at
 CC serine and/or threonine residues. The protein has potential antimicrobial
 CC and antibacterial activity and is useful in screening for antagonists,
 CC agonists and for compounds with antibiotic activity. The proteins are
 CC also useful in determining their role in pathogenesis of infection,
 CC dysfunction and disease and could be used as part of a vaccine and/or
 CC peptide therapy.
 xx

SO Sequence 448 AA;

Query Match: 43.3%; Score 935.5; DB 22; Length 448;
 Best Local Similarity 43.2%; Pred. No. 8.8e-73;

Matches 189; Conservative 86; Mismatches 134; Indels 9; Gaps 4;

OY 6 VAIVGRPNVGSSTIFNRIVGERVSTIEDTPGVTRDRIRYSGEMLTNDFNIIDTGTIEIGD 65
 DB 4 VAIGRTNWKSTLFRLOKPMIAIVSDPTNTRDRIRIGIGEMLRKKAIFIDTGLAKQ 63
 OY 66 APFOTQIRAOAEIADADVIIFPMVNRGELTOSDEMVAOILYKSK-KPVYLAIVKVDNM 124
 DB 64 TPLQOLALQVQAALQSAKAIIFLVLSLOQLNSDDPYAKYVKKKKDKPVLVYKAKENF 123
 OY 125 EMRT---DVFDTSLGFCGPYPISSGSHGIGLDDLDAVYSH---RGEEDPYDEDTIRL 178
 DB 124 NPKAETTLNDYISLGFGRVYISAHAGIGIGLMDLVLKONQLLPNEND--DLAKIRF 181
 OY 179 SIIGRPVNGKSSLVNAILGEDRVIVSNVAGTTRDAIDTEYSYDGDVVLIDTACMRKKG 238
 DB 182 CVIGKPNVGSKSLINQLVKONRVLVNSGSTRDAIDVPLKYNKEKFLIDTACIKRRKG 241
 OY 239 VYESTEKYSVLRLAKIERSNVVLVIADEOGITIEODKRVAGYAHEGKAVVIYVNNKMDT 298
 DB 242 INNCIETASTYIKTKLARSNYLLNVDOSKPSISEDEVIGLQAALIPYILVNNKMDL 301
 OY 299 VEKDSKTMKFEDEVKREFOLDYAOIAFVSAKERTRLTLEPYINEASENKKRVOST 358
 DB 302 VLKNNNTNAVKKMLKLFKHLDFAFVLFSVLKNQRLNTFEOLKIQSOLETKVATPL 361
 OY 359 LNEYVTDATSMNPTPTDGRRLNVEFATOVATEPPTFVVFNVDVLMHFSYKRYLENOIR 418
 DB 362 LNDVIOQAOLNQLPFLKGRKRLQITAVOTKSOIPIHFVLCNDPKYILHFSYARLENKIR 421
 OY 419 AARGECTPIHITAKRN 436
 DB 422 ENFGFNSVPISTLYFSKN 439

Search completed: September 10, 2003, 00:32:44
 Job time : 88 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Comphen Ltd.

OM protein - protein search, using sw model

Run on: September 10, 2003, 00:29:51; Search time 43 Seconds

(without alignments)
975.106 Million cell updates/sec

Title: US-09-815-242-12600

Perfect score: 2213

Sequence: 1 MTKPIVAIVGRPNVCKSTIF.....IKAAFGFGCTPIHIAKRN 436

Scoring table: BIOSUM62
Gapop 10.0, Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0
Maximum DB seq length: 200000000Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database: 1: p1r1:.*
2: p1r2:.*
3: p1r3:.*
4: p1r4:.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2213	100.0	436	2	C89926 hypothetical prote
2	1731	78.2	436	2	A11316 hypothetical prote
3	1724	77.9	436	2	A11688 hypothetical prote
4	1699	76.8	436	2	A69336 probable GTP bindi
5	1658	74.9	437	2	F83854 hypothetical prote
6	1522	68.8	436	2	B95199 phosphoglycerate d
7	1520	68.7	436	2	H98065 GTP-binding protei
8	1218.5	55.1	438	2	C86719 GTP-binding protei
9	1207	46.9	451	2	B97111 GTP binding protei
10	1037	46.1	452	2	A81867 GTP binding protei
11	1021	45.0	442	2	S75645 conserved GTP bindi
12	995	44.5	442	2	H98065 conserved GTP bindi
13	984.5	44.5	448	2	D64236 GTP-binding protei
14	935.5	42.3	448	2	E91050 hypothetical prote
15	923	41.7	503	2	A85895 probable GTP-bind
16	923	41.7	503	2	F65027 probable GTP-bind
17	920	41.1	490	2	AF0821 GTP-binding protei
18	910	41.0	494	2	B82284 GTP-binding protei
19	906.5	40.5	495	2	AC0350 GTP-binding protei
20	896.5	39.1	433	2	A70465 GTP-binding protei
21	864.5	38.9	485	2	G81149 essential GTPase N
22	860.5	38.8	485	2	F81871 probable GTP-bind
23	852	38.5	504	2	F64143 conserved GTPase/GT
24	842.5	38.1	439	2	B72253 conserved GTPase/GT
25	840.5	38.0	493	2	B83171 conserved GTPase/GT
26	838	37.9	588	2	D87454 GTP-binding protei
27	806.5	36.4	463	2	H70504 GTP-binding protei
28	797.5	36.0	483	2	AH3445 GTP-binding protei

30	789	35.7	476	2	A97636 probable GTP-bind
31	789	35.7	476	2	AC2859 GTP-binding protei
32	784	35.4	438	2	F75290 conserved GTP-bind
33	782.5	35.4	461	2	S72953 probable GTP-bind
34	782.5	35.4	462	2	F87080 GTP-binding protei
35	779	35.2	465	2	D82804 GTP-binding protei
36	778.5	35.2	475	2	A95982 GTP-binding protei
37	752	34.0	453	2	C85000 hypothetical GTP-b
38	746.5	33.7	447	2	F97828 conserved GTP-bind
39	737.5	33.3	447	2	H71673 hypothetical prote
40	734.5	33.2	460	2	H81381 probable GTP-bind
41	719.5	32.5	490	2	A81744 GTP-binding protei
42	710.5	32.1	490	2	H71480 probable GTPase/GT
43	706	31.9	458	2	B64624 GTP-binding protei
44	701	31.7	462	2	E71891 GTP-binding protei
45	685.5	31.0	487	2	B86596 GTPase/GTP-binding

ALIGNMENTS

RESULT 1

C89926 hypothetical protein SA1307 [Imported] - Staphylococcus aureus (strain N315)

C:Species: Staphylococcus aureus

C:Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 22-Oct-2001

C:Accession: C89926

R:Kuroda, M.; Onita, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; O

ma, A.; Mizutani-Uji, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K

C.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiratazu, K.

A:Title: Whole genome sequencing of methicillin-resistant Staphylococcus aureus.

A:Reference number: AB9758; MUID:21311952; PMID:11418146

A:Accession: C89926

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-436 <KUR>

A:Cross-references: GB:BA000018; PID:913701274; PIDN:BA842568.1; GSPDB:GN00149

A:Experimental source: strain N315

C:Genetics:

C:Gene: SA1307

C:Superfamily: Mycobacterium leprae probable GTP-binding protein; translation elongat

Query Match	Score	DB 2	Length	436
Best local similarity	100.0%	100.0%	Pred. No. 1e-128	
Matches	436	Conservative	0	Mismatches
			0	Indels
			0	Gaps
			0	
OY	1	MTKPIVAIVGRPNVCKSTIFNRIVGERVSTVEDTPCVTRDRITSSGEMLTDFNITDTCG	60	
DB	1	MTKPIVAIVGRPNVCKSTIFNRIVGERVSTVEDTPCVTRDRITSSGEMLTDFNITDTCG	60	
OY	61	IEIGDAPFQOTIRAOAEIAIDADVILFVNVRGLTQSDENVAOILYKSKPPVLAANK	120	
DB	61	IEIGDAPFQOTIRAOAEIAIDADVILFVNVRGLTQSDENVAOILYKSKPPVLAANK	120	
OY	121	VNNMRTDYDFSLSGEPYISGSHGLGLDLDAVVSHPGEEDPDEDTIRLSI	180	
DB	121	VNNMRTDYDFSLSGEPYISGSHGLGLDLDAVVSHPGEEDPDEDTIRLSI	180	
OY	181	ICRPVNGKSSLVNMLICEDRVIVSNVAGTTRDAIDTETSYDGDVYLIDTAGMRKKNVY	240	
DB	181	ICRPVNGKSSLVNMLICEDRVIVSNVAGTTRDAIDTETSYDGDVYLIDTAGMRKKNVY	240	
OY	241	ESTETSYVLRALKIERSNVVLYIDAEGGIIEDQDRVAGVYHEDGKAVIVVNMKDPVE	300	
DB	241	ESTETSYVLRALKIERSNVVLYIDAEGGIIEDQDRVAGVYHEDGKAVIVVNMKDPVE	300	
OY	301	KOSKTKMKFFEDVKKEQFLDYAOIAVSKKERTRLTLPPIYNASEHKKRVOSSFLN	360	
DB	301	KOSKTKMKFFEDVKKEQFLDYAOIAVSKKERTRLTLPPIYNASEHKKRVOSSFLN	360	
OY	361	EVVTAIISMNPTPTDKGRLNVFATOVAALEPPTFVFNDELHMFYSKRYLENOIRAA	420	
DB	361	EVVTAIISMNPTPTDKGRLNVFATOVAALEPPTFVFNDELHMFYSKRYLENOIRAA	420	

Db 361 EVYTDALISMNPTDGRRLNLFATQVAIEPTFFVFNVDVLMHFSYKRYLENOIRAA 420
 Oy 421 FGEGCTPIHITARRKN 436
 Db 421 FGEGCTPIHITARRKN 436

RESULT 2

hypothetical protein lmo1937 [imported] - listeria monocytogenes (strain EGD-e)
 A:Accession: A11316
 C:Species: Listeria monocytogenes
 C:Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 14-Dec-2001
 R:Classer, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloeker, D.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurge, O.; Entlian, K.D.; Fsihl, H.; Science 294, 849-852, 2001
 A:Authors: Krefte, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maltournam, A.; Ok, C.; Schlueter, T.; Simoes, N.; Tjere, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland, A.; Title: Comparative genomics of Listeria species.
 A:Reference number: AB1077; MUID:21537279; PMID:11679669
 A:Accession: A11316
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-436 <GLA>
 A:Cross-references: GB:NC_003210; PIDN:CAD00015.1; PID:g16411390; GSPDB:GN00177
 A:Experimental source: strain EGD-e
 C:Genetics:
 A:Gene: lmo1937
 C:Superfamily: Mycobacterium leprae probable GTP-binding protein; translation elongation

Query Match 78.2% Score 1731; DB 2: Length 436;
 Best Local Similarity 73.6% Pred. No. 4, 3e-99;
 Matches 320; Conservative 65; Mismatches 50; Indels 0; Gaps 0;

Oy 1 MKRPVAIVGRPNVCKSTIFNRIVGERVSVEDTPGVTRDRISGCEMLTHDFNIIDTCG 60
 Db 1 MAKPVVAIVGRPNVCKSTIFNRIVGERVSVEDTPGVTRDRISGCEMLTHDFNIIDTCG 60
 Oy 61 IEIGDAPFOTOIRAOAEIAIDADVIFPMVNRGCLTOSDEMVAOILYKSKRPVLAANK 120
 Db 61 IDLSDFPLEOIRAOAEIAIDADVIFITNGREGVTDADQVAKILYKSKRPVLAANK 120
 Oy 121 VDNMEKRTDVPYFSLGCEPRPISGSHGLCLDLDVAVSHGEEDEPDEYEDTIRLSI 180
 Db 121 VDNPEMRDIDYFSLGCEPRPISGSHGLCLDLDVAVSHGEEDEPDEYEDTIRLSI 180
 Oy 181 IGRPNVCKSSILNALGCDRIVSVNAGTTRDAIDTETSYDGDVYLIDTAGMRKRGVY 240
 Db 181 IGRPNVCKSSILNALGCDRIVSVNAGTTRDAIDTETSYDGDVYLIDTAGMRKRGVY 240
 Oy 241 ESTEKYSVLRALKAIERSNVLVYIDAEOGIEODKRVAGYAHEOGKAVIYVNNMDPVE 300
 Db 241 ESTEKYSVLRALKAIERSNVLVYIDAEOGIEODKRVAGYAHEOGKAVIYVNNMDPVE 300
 Oy 301 KDSKTKMKFEDEVRKEOPFLDYAOIAVSAKERTRLTLPYINASENHKKRVOSSTLN 360
 Db 301 KDSKTKMKFEDEVRKEOPFLDYAOIAVSAKERTRLTLPYINASENHKKRVOSSTLN 360
 Oy 361 EVYTDALISMNPTDGRRLNLFATQVAIEPTFFVFNVDVLMHFSYKRYLENOIRAA 420
 Db 361 EVYTDALISMNPTDGRRLNLFATQVAIEPTFFVFNVDVLMHFSYKRYLENOIRAA 420
 Oy 421 FGEGCTPIHITARRK 435
 Db 421 FGEGCTPIHITARRK 435

RESULT 3

hypothetical protein homolog lln2051 [imported] - listeria innocua (strain Clp11262)
 C:Species: Listeria innocua
 C:Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 14-Dec-2001

data no good

C:Accession: A11688
 R:Classer, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloeker, D.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurge, O.; Entlian, K.D.; Fsihl, H.; Science 294, 849-852, 2001
 A:Authors: Krefte, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maltournam, A.; Ok, C.; Schlueter, T.; Simoes, N.; Tjere, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland, A.; Title: Comparative genomics of Listeria species.
 A:Reference number: AB1077; MUID:21537279; PMID:11679669
 A:Accession: A11688
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-436 <GLA>
 A:Cross-references: GB:AL592022; PIDN:CAC97281.1; PID:g16414552; GSPDB:GN00178
 A:Experimental source: strain Clp11262
 C:Genetics:
 A:Gene: lln2051
 C:Superfamily: Mycobacterium leprae probable GTP-binding protein; translation elongation

Query Match 77.9% Score 1724; DB 2: Length 436;
 Best Local Similarity 73.1% Pred. No. 1, 2e-98;
 Matches 318; Conservative 67; Mismatches 50; Indels 0; Gaps 0;

Oy 1 MKRPVAIVGRPNVCKSTIFNRIVGERVSVEDTPGVTRDRISGCEMLTHDFNIIDTCG 60
 Db 1 MAKPVVAIVGRPNVCKSTIFNRIVGERVSVEDTPGVTRDRISGCEMLTHDFNIIDTCG 60
 Oy 61 IEIGDAPFOTOIRAOAEIAIDADVIFPMVNRGCLTOSDEMVAOILYKSKRPVLAANK 120
 Db 61 IDLSDFPLEOIRAOAEIAIDADVIFITNGREGVTDADQVAKILYKSKRPVLAANK 120
 Oy 121 VDNMEKRTDVPYFSLGCEPRPISGSHGLCLDLDVAVSHGEEDEPDEYEDTIRLSI 180
 Db 121 VDNPEMRDIDYFSLGCEPRPISGSHGLCLDLDVAVSHGEEDEPDEYEDTIRLSI 180
 Oy 181 IGRPNVCKSSILNALGCDRIVSVNAGTTRDAIDTETSYDGDVYLIDTAGMRKRGVY 240
 Db 181 IGRPNVCKSSILNALGCDRIVSVNAGTTRDAIDTETSYDGDVYLIDTAGMRKRGVY 240
 Oy 241 ESTEKYSVLRALKAIERSNVLVYIDAEOGIEODKRVAGYAHEOGKAVIYVNNMDPVE 300
 Db 241 ESTEKYSVLRALKAIERSNVLVYIDAEOGIEODKRVAGYAHEOGKAVIYVNNMDPVE 300
 Oy 301 KDSKTKMKFEDEVRKEOPFLDYAOIAVSAKERTRLTLPYINASENHKKRVOSSTLN 360
 Db 301 KDSKTKMKFEDEVRKEOPFLDYAOIAVSAKERTRLTLPYINASENHKKRVOSSTLN 360
 Oy 361 EVYTDALISMNPTDGRRLNLFATQVAIEPTFFVFNVDVLMHFSYKRYLENOIRAA 420
 Db 361 EVYTDALISMNPTDGRRLNLFATQVAIEPTFFVFNVDVLMHFSYKRYLENOIRAA 420
 Oy 421 FGEGCTPIHITARRK 435
 Db 421 FGEGCTPIHITARRK 435

RESULT 4

probable GTP binding protein yphc - Bacillus subtilis
 C:Species: Bacillus subtilis
 C:Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 02-Feb-2001
 C:Accession: A69936

R:Kunst, F.; Ogasawara, N.; Mosser, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Ber, C.; Bron, S.; Brouillet, S.; Brusch, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; A.; Ehrlich, S.D.; Emerson, P.T.; Entlian, K.D.; Errington, J.; Fabret, C.; Ferrari, Nature 390, 249-256, 1997
 A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Funai, S.; Galizzi, A.; Galizzi, J.; Harwood, C.R.; Henault, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M.; Koetter, P.; Koningsstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardino, A.; Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Men, Y.; M. Ogasawa, K.; Ogata, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portee, Rieger, M.; Rivolta, C.; Roche, E.; Roche, B.; Rose, M.; Sadale, Y.; Sato, T.; Scani, A.; Authors: Schleich, S.; Schroeter, R.; Scollone, F.; Sekiguchi, J.; Sekowska, A.; Se

akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpstra, P.; Tognoni, A.; Tosato, V.; Uchiyama, T.; Winters, P.; Wipac, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, K.
 A:Authors: Yoshikawa, H.F.; Zumbstein, E.; Yoshikawa, H.; Danchin, A.
 A:Title: The complete genome sequence of the Gram-positive bacterium *Bacillus subtilis*.
 A:Reference number: A69580; MID:98044033; PMID:9384377
 A:Accession: A69936
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1436 <KUN>
 A:Cross-references: GB:299115; GB:AL009126; MID:92634478; PIDN:CAB14200.1; PID:92634702
 A:Experimental source: strain 168
 C:Genetics:
 A:Gene: yphC
 C:Superfamily: Mycobacterium leprae probable GTP-binding protein; translation elongation
 C:Keywords: duplication; GTP binding; nucleotide binding; P-loop
 F:4-122/Domain: translation elongation factor Tu homology <ET1>
 F:10-17/Region: nucleotide-binding motif A (P-loop)
 F:119-112/Region: GTP-binding NKX motif
 F:1145-117/Domain: translation elongation factor Tu homology <ET2>
 F:1175-297/Domain: GTP-binding SAK/L motif status atypical
 F:182-189/Region: translation elongation factor Tu homology <ET2>
 F:182-297/Region: GTP-binding NKX motif
 F:294-331/Region: GTP-binding SAK/L motif
 F:329-331/Region: GTP-binding SAK/L motif
 Query Match 76.84; Score 1699; DB 2; Length 436;
 Best Local Similarity 73.64; Pred. No. 46-97; Indels 0; Gaps 0;
 Matches 320; Conservative 57; Mismatches 58; Indels 0; Gaps 0;
 QY 1 MTKPIVAIVGRPNVNGKSTIFNRIVGERVSIEDPTGVTDRIRYSSEGLTHDFNIIDTGC 60
 DB 1 MGRVVAIVGRPNVNGKSTIFNRIVGERVSIEDPTGVTDRIRYSSEGLTHDFNIIDTGC 60
 QY 61 IEIGDAPFOQIRAGAEIADADVIFPMVNRREGLTOSDEKVAQILYKSKRPVLA VNK 120
 DB 61 IDICDEPFLAQIRQAEIADADVIFPMVNRREGLTOSDEKVAQILYKSKRPVLA VNK 120
 QY 121 VDNMEMRTDYYDFYSLGCEPYPISGSHGLGDLDAVSHSGEEDDPYDEDTIRLSI 180
 DB 121 LDNTKEMANITDPTSLGCEPYPISGSHGLGDLDAVSHSGEEDDPYDEDTIRLSI 180
 QY 181 IGRPNVNGKSSLVNAMLGEERVIYSNVAGTTRDAVDTSTYNOQEVYIDTGMKRGKY 240
 DB 181 IGRPNVNGKSSLVNAMLGEERVIYSNVAGTTRDAVDTSTYNOQEVYIDTGMKRGKY 240
 QY 241 ESTEKYSVLRALKAIERSNVLVVDAEQGIIEODKRVAGVAHEGKAAVIVVNMKDPVE 300
 DB 241 ETTEKYSVLRALKAIERSNVLVVDAEQGIIEODKRVAGVAHEGKAAVIVVNMKDAVD 300
 QY 301 KDSITMKKFEDEVRKEFOFLDYAOIAFVSAKERTLRILFPIINASENHKKRVQSTLN 360
 DB 301 KDSITMKKFEDEVRKEFOFLDYAOIAFVSAKERTLRILFPIINASENHKKRVQSTLN 360
 QY 361 EYVDAISMNPTPTDKGRRLNVFATQVAIEPPTFVFNVDVLMHSEYKRYLENOIRAA 420
 DB 361 DVIMDAVAMNPTPTDHNKGRRLNVFATQVAIEPPTFVFNVDVLMHSEYKRYLENOIRAA 420
 QY 421 FGEGCPPIHTIARRK 435
 DB 421 FGEGCPPIHTIARRK 435
 RESULT 5
 F83854
 hypochelical protein BHI638 [imported] - *Bacillus halodurans* (strain C-125)
 C:Species: *Bacillus halodurans*
 C:Date: 01-Dec-2000 #sequence, revision 01-Dec-2000 #text_change 15-Jun-2001
 C:Accession: F83854
 R:Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; Hieda
 Nucleic Acids Res. 28, 4317-4331, 2000
 A:Title: Complete genome sequence of the alkaliphilic bacterium *Bacillus halodurans* and
 A:Reference number: A83550; MID:20512582; PMID:11058132
 A:Accession: F83854
 A:Status: preliminary

A:Molecule type: DNA
 A:Residues: 1437 <STO>
 A:Cross-references: GB:A000512; GB:BA000004; MID:910174030; PIDN:BA053357.1; GSPDB:C
 A:Experimental source: strain C-125
 C:Genetics:
 A:Gene: BHI638
 C:Superfamily: Mycobacterium leprae probable GTP-binding protein; translation elongation
 Query Match 74.98; Score 1658; DB 2; Length 437;
 Best Local Similarity 70.96; Pred. No. 1,36-94;
 Matches 309; Conservative 67; Mismatches 60; Indels 0; Gaps 0;
 QY 1 MTKPIVAIVGRPNVNGKSTIFNRIVGERVSIEDPTGVTDRIRYSSEGLTHDFNIIDTGC 60
 DB 1 MTKPIVAIVGRPNVNGKSTIFNRIVGERVSIEDPTGVTDRIRYSSEGLTHDFNIIDTGC 60
 QY 61 IEIGDAPFOQIRAGAEIADADVIFPMVNRREGLTOSDEKVAQILYKSKRPVLA VNK 120
 DB 61 IEIGDAPFOQIRAGAEIADADVIFPMVNRREGLTOSDEKVAQILYKSKRPVLA VNK 120
 QY 121 VDNMEMRTDYYDFYSLGCEPYPISGSHGLGDLDAVSHSGEEDDPYDEDTIRLSI 180
 DB 121 IDHPDMOELTEFEYSLGCEPYPISGSHGLGDLDAVSHSGEEDDPYDEDTIRLSI 180
 QY 181 IGRPNVNGKSSLVNAMLGEERVIYSNVAGTTRDAVDTSTYNOQEVYIDTGMKRGKY 240
 DB 181 IGRPNVNGKSSLVNAMLGEERVIYSNVAGTTRDAVDTSTYNOQEVYIDTGMKRGKY 240
 QY 241 ESTEKYSVLRALKAIERSNVLVVDAEQGIIEODKRVAGVAHEGKAAVIVVNMKDPVE 300
 DB 241 ESTEKYSVLRALKAIERSNVLVVDAEQGIIEODKRVAGVAHEGKAAVIVVNMKDAVE 300
 QY 301 KDSITMKKFEDEVRKEFOFLDYAOIAFVSAKERTLRILFPIINASENHKKRVQSTLN 360
 DB 301 KDSITMKKFEDEVRKEFOFLDYAOIAFVSAKERTLRILFPIINASENHKKRVQSTLN 360
 QY 361 EYVDAISMNPTPTDKGRRLNVFATQVAIEPPTFVFNVDVLMHSEYKRYLENOIRAA 420
 DB 361 DVIMDAVAMNPTPTDHNKGRRLNVFATQVAIEPPTFVFNVDVLMHSEYKRYLENOIRAA 420
 QY 421 FGEGCPPIHTIARRK 436
 DB 421 FGEGCPPIHTIARRK 436
 RESULT 6
 B95199
 phosphoglycerate dehydrogenase-related protein [imported] - *Streptococcus pneumoniae*
 C:Species: *Streptococcus pneumoniae*
 C:Date: 03-Aug-2001 #sequence, revision 03-Aug-2001 #text_change 24-Aug-2001
 C:Accession: B95199
 R:Petelin, H.; Nelson, K.E.; Paulsen, I.T.; Eisen, J.A.; Read, T.D.; Peterson, S.; H
 nson, T.J.; Hickey, E.K.; Holt, I.E.
 Science 293, 498-506, 2001
 A:Authors: Loftus, B.J.; Yang, F.; Smith, H.O.; Venter, J.C.; Dougherty, B.A.; Morris
 A:Title: Complete Genome Sequence of a Virulent Isolate of *Streptococcus pneumoniae*.
 A:Reference number: A95000; MID:21357203; PMID:11465916
 A:Accession: B95199
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1436 <KUN>
 A:Cross-references: GB:A0005672; PIDN:AAK75187.1; PID:914973204; GSPDB:GN00164; TIGR:
 C:Genetics:
 A:Gene: SPI709
 C:Superfamily: Mycobacterium leprae probable GTP-binding protein; translation elongation
 Query Match 68.88; Score 1522; DB 2; Length 436;
 Best Local Similarity 66.28; Pred. No. 36-86;
 Matches 290; Conservative 73; Mismatches 69; Indels 6; Gaps 3;
 QY 1 MTKPIVAIVGRPNVNGKSTIFNRIVGERVSIEDPTGVTDRIRYSSEGLTHDFNIIDTGC 60

F:295-298/Region: GTP-binding NKXD motif
F:330-332/Region: GTP-binding SAK/L motif #status atypical

Query Match 46.1% Score 1021, DB 2; Length 452;
Best Local Similarity 45.9%: Pred. No. 2e-55;
Matches 200; Conservative 94; Mismatches 138; Indels 4; Gaps 4;

QY 1 MTKPIVAIGRPVNGKSTIFNRIGERSIVEDTGTTRDRIRYSSGEMLTHTDNIIDTGG 60
DB 1 MSPLVAIGRPVNGKSTIFNRVLAGNQAIVHDGTTTRDRIRYAPFMRDRDQVVDIG 60
QY 61 IELG-DAPFOQIRAOEIAIDEDAVIIFMNVNREGTOSDENVAQILYKSKRPVLA 119
DB 61 LVFNDSFELPEIRDOALALAEKAAIFVVDGQGGTASDEEIAQMLKQOSVPLA 120
QY 120 KVDNMEM-RTDYDFYSLGCEPPYPIGSGHGLGDLDAVNSHFEEDEDPYDEDTIR 178
DB 121 KCESPDQAIQAAEFMHLGLCEPYFMSAHGSGDLDLLETPAPQEEP-EDEEIKY 179
QY 179 STIGRPVNGKSLVAALIGEDRVIVSNVAGTTTBAIDTEXSYDGODVLIIDTGMKRGK 238
DB 180 AIVGRPNVNGKSSLLAALGEGRAIVSPISGTTTBAIDMVYERKQKRLIDTIGIRKK 239
QY 239 VESTERKYSVLAKAERSNVVLVIDAEGGTEEDKRVAGYAHGOKAVIVNKKMDT 298
DB 240 VDYGAFFGGINRAFAIRADVLLVLDVLDGVTEDDLKLAGRIEDGKRAVVLVINKMDA 299
QY 299 VERKSKTKMKFEDEKREFOLDYAOIAFVSAKERTRLTLPYINAESENHKKRVOSST 358
DB 300 VERKSDYITIEHREQLMARLYFMDNAEMIFVSAQICGLAVQILDCVDAEHRRTYAV 359
QY 359 LNDVYTAISMNPPTDK-GRRLVAFYATOVAIPEPFVFNVDVLMHSYKRYLENOI 417
DB 360 IREVLBEANVSHSPRTKQKQKIVYGTOSTOPPAIALVNDPKNRNDYRYTEKOF 419
QY 418 RAASFECTPIHIIAR 433
DB 420 RKQCGFGSPILRFLWR 435

RESULT 12
882899
conserved hypothetical ATP/GTP-binding protein U0383 (imported) - Ureaplasma urealyticum
C:Species: Ureaplasma urealyticum
C>Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 02-Sep-2000
C:Accession: B82899
R:Glass, J.I.; Lefkowitz, E.J.; Glass, J.S.; Helner, C.R.; Chen, E.Y.; Cassell, G.H.
submitted to GenBank, February 2000
A:Description: The complete sequence of Ureaplasma urealyticum: Alternate views of a mit
A:Reference number: A82870
A:Accession: B82899
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-442 <GIA>
A:Cross-references: GB:A8002135; GB:AF222894; NID:96899358; PIDN:AAF30793.1; GSPDB:GN001
A:Experimental source: setovar 3; biovar 1
C:Genetics:
A:Gene: U0383
A:Genetic code: SGC3
C:Superfamily: Mycobacterium leprae probable GTP-binding protein: translation elongator
F:175-296/Domain: translation elongation factor tu homology <ENU>

Query Match 45.0% Score 995, DB 2; Length 442;
Best Local Similarity 45.4%: Pred. No. 7.6e-54;
Matches 137; Conservative 93; Mismatches 140; Indels 4; Gaps 4;

QY 6 VAIVGRPNVNGKSTIFNRIGERSIVEDTGTTRDRIRYSSGEMLTHTDNIIDTGG 65
DB 4 IAIIVKPNVNGKSSLPNRLMRKRSIVDDGCVTRDRIRYDIGNMLTSPMLDTGGIISK 63
QY 66 AFPOQIRAOEIAIDEDAVIIFMNVNREGTOSDENVAQIL-YKSK-RPVVLA 123
DB 64 DTVOONINQVFLPAINENATITIFVSAKDGINDDKIAIMLAKKAKKIIIVINKIES 123

QY 124 MEKRTDYDFYSLGCEPPYPIGSGHGLGDLDAVNSHFEEDEDPYDEDTIRLSICR 183
DB 124 EKYLLNGELYSFGFKFKFIRSAEHGICGDLDELVDN-PIONNKDERKFCRIR 182
QY 184 PNVKSSLVNAILIGEDRVIVSNVAGTTTBAIDTEXSYDGODVLIIDTGMKRGKYE 243
DB 183 PNVKSSLVNAILIGEDRVIVSNVAGTTTBAIDTEXSYDGODVLIIDTGMKRGKYE 242
QY 244 EKYSLVRLAKAERSNVVLVIDAEGGTEEDKRVAGYAHGOKAVIVNKKMDT-VERD 302
DB 243 EKYAVLTKKALRSOLILVLDGSEPREDEDEVGGLATIANPTIIVNKKMDINIKK 302
QY 303 SKTKKFEDEKREFOLDYAOIAFVSAKERTRLTLPYINAESENHKKRVOSSTLE 362
DB 303 SHTEMVKKOIRSOFFYLSMAPIVVSAIDNKKRIHTIFEALIEFVREOAMKRIATSL 362
QY 363 VTDAISMNPPTDKGRRLVAFYATOVAIPEPFVFNVDVLMHSYKRYLENOIAR 422
DB 363 VIANANQEPPEPKGGRISISIVYQSOIPTFVLKCNRPFLHFVARIENIKRG 422
QY 423 FEGTPIHIIARKRN 436
DB 423 FDSVPTITLVWQDN 436

RESULT 13
H90522
GTP-binding protein (imported) - Mycoplasma pulmonis (strain UAB CTIP)
C:Species: Mycoplasma pulmonis
C>Date: 24-May-2001 #sequence_revision 24-May-2001 #text_change 03-Aug-2001
C:Accession: H90522
R:Chambaud, I.; Hellwig, R.; Ferris, S.; Barbe, V.; Samson, D.; Gallsen, F.; Moszer, Nucleic Acids Res. 29, 2145-2153, 2001
A:Title: The complete genome sequence of the murine respiratory pathogen Mycoplasma p
A:Reference number: A99512; NUID:21267165; PMID:11353084
A:Accession: H90522
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-435 <KUR>
A:Cross-references: GB:A1445566; PID:G14089501; PIDN:CAC13261.1; GSPDB:GN00153
A:Experimental source: strain UAB CTIP
C:Genetics:
A:Gene: MYPU_0880
A:Genetic code: SGC3
C:Superfamily: Mycobacterium leprae probable GTP-binding protein: translation elongat

Query Match 44.5% Score 984.5, DB 2; Length 435;
Best Local Similarity 43.5%: Pred. No. 3.3e-53;
Matches 187; Conservative 108; Mismatches 130; Indels 5; Gaps 3;

QY 5 IVAIVGRPNVNGKSTIFNRIGERSIVEDTGTTRDRIRYSSGEMLTHTDNIIDTGG 64
DB 9 IVAIVGRPNVNGKSTIFNRVLAGNQAIVHDGTTTRDRIRYAPFMRDRDQVVDIG 68
QY 65 DAPFOQIRAOEIAIDEDAVIIFMNVNREGTOSDENVAQILYKSKRPVLA 124
DB 65 DAPFOQIRAOEIAIDEDAVIIFMNVNREGTOSDENVAQILYKSKRPVLA 123
QY 125 EMRTDYDFYSLGCEPPYPIGSGHGLGDLDAVNSHFEEDEDPYDEDTIRLSICR 184
DB 128 -AKNEDSYSLGCEPPYPIGSGHGLGDLDAVNSHFEEDEDPYDEDTIRLSICR 184
QY 185 NVKSSLVNAILIGEDRVIVSNVAGTTTBAIDTEXSYDGODVLIIDTGMKRGKYE 244
DB 185 NVKSSLVNAILIGEDRVIVSNVAGTTTBAIDTEXSYDGODVLIIDTGMKRGKYE 244
QY 245 KYSVLAKAERSNVVLVIDAEGGTEEDKRVAGYAHGOKAVIVNKKMDTVERD 302
DB 245 KYSVLAKAERSNVVLVIDAEGGTEEDKRVAGYAHGOKAVIVNKKMDTVERD 304
QY 305 TMKFEDEKREFOLDYAOIAFVSAKERTRLTLPYINAESENHKKRVOSSTLE 364
DB 305 TMKFEDEKREFOLDYAOIAFVSAKERTRLTLPYINAESENHKKRVOSSTLE 364

DR Interpro: IPR002917; MMR_HSR1.
 DR Interpro: IPR005225; Small_GTP.
 DR Pfam: PF01926; MMR_HSR1. 1.
 DR PRINTS: PR00326; GTP1_OBG.
 DR TIGRFAMs: TIGR00650; MG442; 2.
 DR TIGRFAMs: TIGR00231; small_GTP; 2.
 DR GTP-binding: Repeat: Complete proteome.
 KM GTP-binding: Repeat: Complete proteome.
 FT NP_BIND 10 17 GTP 1 (POTENTIAL).
 FT NP_BIND 57 61 GTP 1 (POTENTIAL).
 FT NP_BIND 119 122 GTP 1 (POTENTIAL).
 FT NP_BIND 182 189 GTP 2 (POTENTIAL).
 FT NP_BIND 229 233 GTP 2 (POTENTIAL).
 FT NP_BIND 294 297 GTP 2 (POTENTIAL).
 SQ SEQUENCE 436 AA: 48979 MW: 9544281862F229D3 CRC64;

Query Match 100.0%; Score 2213; DB 1; Length 436;
 Best Local Similarity 100.0%; Pred. No. 2, 7e-128;
 Matches 436; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 MKRPVIAVGRPNVSKSTIFNRIVGERVSIYEDTPGVTRDRITVSSGEMLTHTDNIIDTGC 60
 DB 1 MKRPVIAVGRPNVSKSTIFNRIVGERVSIYEDTPGVTRDRITVSSGEMLTHTDNIIDTGC 60
 OY 61 IETGAPFOTOIRAQAEIAIDADVIFPMVNRGLTQSDENVAQILYKSKRPVLAANK 120
 DB 61 IETGAPFOTOIRAQAEIAIDADVIFPMVNRGLTQSDENVAQILYKSKRPVLAANK 120
 OY 121 VONMEMKRTDYDFYSLGCEPPIPSGSHGLGDLDAVYSHGEEDEDPYEDDTIRLSI 180
 DB 121 VONMEMKRTDYDFYSLGCEPPIPSGSHGLGDLDAVYSHGEEDEDPYEDDTIRLSI 180
 OY 181 IGRPNVGSLLVNALGSDRYVSNVAGTTRDAIDTETSYDGDYVLIIDTACRRKRGKY 240
 DB 181 IGRPNVGSLLVNALGSDRYVSNVAGTTRDAIDTETSYDGDYVLIIDTACRRKRGKY 240
 OY 241 ESTEKYSVLRAKALERSNVLVYIDAEGCIIEEDKRVAGYAHDOGRAVYVNNKMDYVE 300
 DB 241 ESTEKYSVLRAKALERSNVLVYIDAEGCIIEEDKRVAGYAHDOGRAVYVNNKMDYVE 300
 OY 301 KOSKTKKFEDEVREKEFQFLDYAOIAFVSAKERTRLTLPYINAESENHKKRVOSTLN 360
 DB 301 KOSKTKKFEDEVREKEFQFLDYAOIAFVSAKERTRLTLPYINAESENHKKRVOSTLN 360
 OY 361 EYVTDASNPPTDKGRLLVFAVATOAIEPPFFVYFVNDVDELHMFYSKRYLENDIRAA 420
 DB 361 EYVTDASNPPTDKGRLLVFAVATOAIEPPFFVYFVNDVDELHMFYSKRYLENDIRAA 420
 OY 421 FGEGTPIHIIARKRN 436
 DB 421 FGEGTPIHIIARKRN 436

RESULT 2
 ENCA_LISTO STANDARD: PRT; 436 AA.
 ID ENCA_LISTO STANDARD: PRT; 436 AA.
 AC 08Y5W8;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DE Probable GTP-binding protein enca.
 GN ENCA OR LMO1937.
 OS Listeria monocytogenes.
 OC Bacteria; Firmicutes; Bacillales; Listeriaceae; Listeria.
 OX NCBI_TaxID=1639;
 RN SEQUENCE FROM N.A.
 RC STRAIN=Egd-e / Serovar 1/2a;
 RX MEDLINE=21537279; PubMed=11679669;
 RA Gasser P., Frangeul L., Buchrieser C., Rusniok C., Amend A.,
 RA Aguero P., Berche P., Bloecher H., Brandt P., Chakraborty T.,
 RA Charlit A., Chetoui F., Couve E., de Darvar A., Dehoux P.,
 RA Domann E., Dominguez-Bernal G., Duchaud E., Durant L., Dussutget O.,
 RA Estlin K.-D., Fainl H., Garcia-del Portillo F., Garrido P.,

RA Gautier L., Goebel W., Gomez-Lopez N., Hain T., Hauf J., Jackson D.,
 RA Jones L.-M., Kaerat U., Kretz J., Kuhn M., Kunst F., Kutzkat G.,
 RA Madueno E., Maitounam A., Mata Vicente J., Ng E., Nedjari H.,
 RA Novales E., Novella S., de Pablos B., Perez-Diaz J.-C., Purcell R.,
 RA Remmel B., Rose M., Schlueter T., Simoes N., Tierrez A.,
 RA Vazquez-Boland J.-A., Voss H., Wehland J., Cossart P.,
 RA "Comparative genomics of *Listeria* species."
 FL Science 294:845-852(2001).
 CC -I- SIMILARITY: BELONGS TO THE ERA/TRME FAMILY OF GTP-BINDING
 CC PROTEINS. ENCA SOBIFAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib.ch).

DR EMBL: AL591981; CAD00015.1; -.
 DR PIR: A1316; A1316.
 DR ListList: LMO1937; -.
 DR HAMAP: MF_00195; -; 1.
 DR Interpro: IPR005289; GTP-binding_dom.
 DR Interpro: IPR006073; GTP_OBG.
 DR Interpro: IPR002917; MMR_HSR1.
 DR Interpro: IPR005225; Small_GTP.
 DR Pfam: PF01926; MMR_HSR1. 1.
 DR PRINTS: PR00326; GTP1_OBG.
 DR TIGRFAMs: TIGR00650; MG442; 2.
 DR TIGRFAMs: TIGR00231; small_GTP; 2.
 DR GTP-binding: Repeat: Complete proteome.
 KM GTP-binding: Repeat: Complete proteome.
 FT NP_BIND 10 17 GTP 1 (POTENTIAL).
 FT NP_BIND 57 61 GTP 1 (POTENTIAL).
 FT NP_BIND 119 122 GTP 1 (POTENTIAL).
 FT NP_BIND 182 189 GTP 2 (POTENTIAL).
 FT NP_BIND 229 233 GTP 2 (POTENTIAL).
 FT NP_BIND 294 297 GTP 2 (POTENTIAL).
 SQ SEQUENCE 436 AA: 49144 MW: 58435C8701A656A3 CRC64;

Query Match 78.2%; Score 1731; DB 1; Length 436;
 Best Local Similarity 73.6%; Pred. No. 7, 6e-99;
 Matches 320; Conservative 65; Mismatches 50; Indels 0; Gaps 0;

OY 1 MKRPVIAVGRPNVSKSTIFNRIVGERVSIYEDTPGVTRDRITVSSGEMLTHTDNIIDTGC 60
 DB 1 MKRPVIAVGRPNVSKSTIFNRIVGERVSIYEDTPGVTRDRITVSSGEMLTHTDNIIDTGC 60
 OY 61 IETGAPFOTOIRAQAEIAIDADVIFPMVNRGLTQSDENVAQILYKSKRPVLAANK 120
 DB 61 IETGAPFOTOIRAQAEIAIDADVIFPMVNRGLTQSDENVAQILYKSKRPVLAANK 120
 OY 121 VONMEMKRTDYDFYSLGCEPPIPSGSHGLGDLDAVYSHGEEDEDPYEDDTIRLSI 180
 DB 121 VONMEMKRTDYDFYSLGCEPPIPSGSHGLGDLDAVYSHGEEDEDPYEDDTIRLSI 180
 OY 181 IGRPNVGSLLVNALGSDRYVSNVAGTTRDAIDTETSYDGDYVLIIDTACRRKRGKY 240
 DB 181 IGRPNVGSLLVNALGSDRYVSNVAGTTRDAIDTETSYDGDYVLIIDTACRRKRGKY 240
 OY 241 ESTEKYSVLRAKALERSNVLVYIDAEGCIIEEDKRVAGYAHDOGRAVYVNNKMDYVE 300
 DB 241 ESTEKYSVLRAKALERSNVLVYIDAEGCIIEEDKRVAGYAHDOGRAVYVNNKMDYVE 300
 OY 301 KOSKTKKFEDEVREKEFQFLDYAOIAFVSAKERTRLTLPYINAESENHKKRVOSTLN 360
 DB 301 KOSKTKKFEDEVREKEFQFLDYAOIAFVSAKERTRLTLPYINAESENHKKRVOSTLN 360
 OY 361 EYVTDASNPPTDKGRLLVFAVATOAIEPPFFVYFVNDVDELHMFYSKRYLENDIRAA 420
 DB 361 EYVTDASNPPTDKGRLLVFAVATOAIEPPFFVYFVNDVDELHMFYSKRYLENDIRAA 420
 OY 421 FGEGTPIHIIARKRN 435

Db 1 |||||:||||
421 FPFEGTPIRVIAKR 435

RESULT 3

ENGA_LISIN STANDARD: PRT: 436 AA.
ID 092A71:
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DE Probable GTP-binding protein engA.
GN ENGA OR LIN2051.
OS *Listeria innocua*.
OC Bacteria; Firmicutes; Bacillales; Listeriaceae; Listeria.
OX NCBI_TaxID=1642;

RP SEQUENCE FROM N.A.
RC STRAIN-CLIP 11262 / Serovar 5a;
RX MEDLINE=21537279; PubMed=11679669;
RA Glaser P., Frangoul L., Buchrieser C., Rusnak C., Amend A.,
RA Baquero F., Berche P., Bloeker H., Brandt P., Chakraborty T.,
RA Chablot A., Chetoui F., Couve E., de Daruvar A., Dehoux P.,
RA Domann E., Dominguez-Bernal G., Duchaud E., Durant L., Dussurget O.,
RA Entian K.-D., Fajal H., Garcia-del Portillo F., Garrido P.,
RA Gautier L., Goebel W., Gomez-Lopez N., Hain T., Haut J., Jackson D.,
RA Jones L.-M., Kaser U., Kreft J., Kuhn M., Kunst F., Kurapatk G.,
RA Madueno E., Maitournam A., Mata Vicente J., Ng E., Nejdari H.,
RA Nordstedt G., Novelle S., de Pablos B., Perez-Diaz J.-C., Purcell R.,
RA Rammel B., Rose M., Schluter T., Simoes N., Tlherre P.,
RA Vazquez-Boland J.-A., Voss H., Weiland J., Cossart P.;
*Comparative genomics of *Listeria* species.*;
RT Science 294:849-852(2001).
RL -1 SIMILARITY: BELONGS TO THE ENA/TRME FAMILY OF GTP-BINDING
PROTEINS. ENGA SUBFAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC EMBL: AL596170; CAC97281.1;
DR PIR: A1688; A1688.
DR HAMAP: MF_00195; -.
DR InterPro: IPR005289; GTP-binding_dom.
DR InterPro: IPR006073; GTP1_OBG.
DR InterPro: IPR002917; MHR_HSR1.
DR InterPro: IPR005225; Small_GTP.
DR Pfam: PF01926; MHR_HSR1; 1.
DR PRINTS: PR00316; GTP1_OBG.
DR TIGRFAMs: TIGR00650; MG442; 2.
DR TRIGRAMs: TIGR00231; small_GTP; 2.
KW GTP-binding: Repeat: Complete proteome.
FT NP_BIND 10 GTP 1 (POTENTIAL).
FT NP_BIND 57 GTP 1 (POTENTIAL).
FT NP_BIND 119 122 GTP 2 (POTENTIAL).
FT NP_BIND 182 189 GTP 2 (POTENTIAL).
FT NP_BIND 229 233 GTP 2 (POTENTIAL).
FT NP_BIND 294 297 GTP 2 (POTENTIAL).
SQ SEQUENCE 436 AA: 49144 MW: 14435C1970B86C8 CRC64;

Query Match 77.9% Score 1724; DB 1; Length 436;
Best Local Similarity 73.1%; Pred. No. 2e-98;

Matches 318; Conservative 67; Mismatches 50; Indels 0; Gaps 0;

OY 1 MKRPVAVGRPNNGKSTIRNRYGERSVETDPTVDRIRSSSEMLTHFNITDGG 60
DB 1 MAKPVAVGRPNNGKSTIRNRYGERSVETDPTVDRIRSSSEMLTHFNITDGG 60

OY 61 IEIGAPFOTOIRAQAEIAIDADYIIFMVNVRGLTQSDENVAOILYKSKRPVLAVK 120
Db 61 IDLSDEPFLEQIRAOAEIAIDADYIIFTVNGREGVADQYAKILYKSKRPVLAVK 120
OY 121 VONHEKRTVYDYSISGFGEPPYISGSHGLGDLDAVYSHGEEDEEDPYEDTRLSI 180
Db 121 VONHEKRTVYDYSISGFGEPPYISGSHGLGDLDAVYSHGEEDEEDPYEDTRLSI 180
OY 181 IGRPNVKGSSLYNALIGEDRYVSNVACTTDAIDTEYSTDGDYVLIIDTACRRKKGY 240
Db 181 IGRPNVKGSSLYNALIGEDRYVSNVACTTDAIDTEYSTDGDYVLIIDTACRRKKGY 240
OY 241 ESTEKSYVALKAIERSNVLVYDAECGITEEDKRVAGYAEQCAVYVNNKMDTVE 300
Db 241 ESTEKSYVALKAIERSNVLVYDAECGITEEDKRVAGYAEQCAVYVNNKMDTVE 300
OY 301 KOSTYKRRFEDREKREFDLOAQAIPVSAEERTLTLFPYINSENNKRRVOSTLN 360
Db 301 KOSTYKRRFEDREKREFDLOAQAIPVSAEERTLTLFPYINSENNKRRVOSTLN 360
OY 361 EVYTDASNNPTDGRRLNVFATOVAVIEPTFVVFVNDVLMHSTYKRYLENOIRAA 420
Db 361 DVISDAVAMNPSPMDKGRKILFYTVQAVAPPTFVVFVNDVLMHSTYKRYLENOIRAA 420
OY 421 FPFEGTPIRVIAKR 435
Db 421 FPFEGTPIRVIAKR 435

RESULT 4

ENGA_BACSU STANDARD: PRT: 436 AA.
ID 092A71:
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Probable GTP-binding protein engA.
GN ENGA
OS *Bacillus subtilis*.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=1423;
RN [1]

RP SEQUENCE FROM N.A.
RC STRAIN-168 / Marburg;
RX MEDLINE=96349105; PubMed=8760912;
RA Sorokin A.V., Azevedo V., Zumstein E., Galleron N., Ehlich S.D.,
RA Serror P.;
FT the sera and Kog loci cloned in a yeast artificial chromosome.*;
RL Microbiology 142:2005-2016(1996).
RN [2]

RP SEQUENCE FROM N.A.
RC STRAIN-168;
RX MEDLINE=9384377; PubMed=9384377;
RA Kunst F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,
RA Azevedo V., Berrtero M.G., Bessieres P., Bolotin A., Borchart S.,
RA Borisov R., Bourcier L., Brans A., Braun M., Bridgell S.C., Bron S.,
RA Brouillet S., Brusch C.V., Caldwell B., Capuano V., Carter N.M.,
RA Choi S.K., Codani J.J., Connerton I.F., Cummings N.J., Daniel R.A.,
RA Denicot F., Devine K.M., Dusterhoft A., Ehlich S.D., Emerson P.A.,
RA Entian K.D., Errington J., Fader C., Ferrari E., Foulger D.,
RA Fritz C., Fujita M., Fujita Y., Fuma S., Galizzi A., Galleron N.,
RA Gilm S.Y., Glaser P., Goffeau A., Golightly E.J., Grandi G.,
RA Gulispi G., Guy B.J., Haga K., Halech J., Harwood C.R., Henaut A.,
RA Hilbert H., Holstappel S., Hosono S., Hullo M.F., Itaya M., Jones L.,
RA Joris B., Karamata D., Kasahara Y., Kleier-Blanchard M., Klein C.,
RA Kobayashi Y., Koeltter P., Koningsstein G., Krogh S., Kumano M.,
RA Kurita K., Lapidos A., Lardinois S., Lauber J., Lazarevic V.,
RA Lee S.M., Levine A., Liu H., Masuda S., Meisel C., Medigue C.,
RA Medina N., Meliadi R.F., Mizuno M., Moesti D., Nakai S., Noback M.,
RA Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B., Park S.H.,
RA Paro V., Pohl T.M., Portetelle D., Porvolik S., Prescott A.M.,
RA Prescan E., Pujic P., Purnelle B., Rapoport G., Rey M., Reynolds S.,

```

RA Rieger M., Rivolta C., Rocha E., Roche B., Rose M., Sadleir Y.,
RA Sato T., Scanlan E., Schleich S., Schroeter R., Scottone F.,
RA Sekiyuchi J., Sekowska A., Sefor S.J., Serror P., Shin B.S., Soldo B.,
RA Scrobin A., Taccori E., Takagi T., Takahashi H., Takemaru K.,
RA Takeuchi M., Tamakoshi A., Tanaka T., Terpiltra P., Tognoni A.,
RA Toseko V., Uchiyama S., Vandenbol M., Vannier F., Vassarotti A.,
RA Vardi A., Wambutt R., Medler E., Medler H., Melzenegger T.,
RA Winters P., Wipac A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,
RA Yoshida K., Yoshikawa H.F., Zumbstein E., Yoshikawa H., Zupchin A.;
RT "The complete genome sequence of the Gram-positive bacterium Bacillus
RT subtilis".
RL Nature 390:249-256(1997).
RN
RN SEQUENCE OF 185-436 FROM N.A.
RP
RP STRAIN-168;
RX MEDLINE-96011379; PubMed-7592341;
RA Morbidoni H.R., de Mendoza D., Cronan J.E. Jr.;
RT "Synthesis of an-glycerol 3-phosphate, a key precursor of membrane
RT lipids, in Bacillus subtilis".
RL J. Bacteriol. 177:5899-5905(1995).
CC -1- SIMILARITY: BELONGS TO THE ERA/TIME FAMILY OF GTP-BINDING
CC PROTEINS. ENCA SUBFAMILY.
CC -1- CAUTION: Ref.3 sequence differs from that shown due to a
CC frameshift.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL: L47648; AAC3966.1;
DR EMBL: Z59115; CAB14200.1;
DR EMBL: U32164; -; NOT_ANNOTATED_CDS.
DR PIR: A69936; A69936.
DR Subtilist; BG11443; enca.
DR HAMAP: MF_00195; -; 1.
DR InterPro: IPR005289; GTP-binding_dom.
DR InterPro: IPR006073; GTP1_OBG.
DR InterPro: IPR002917; MMR_HSR1.
DR InterPro: IPR005225; Small_GTP.
DR Pfam: PF01926; MMR_HSR1; 1.
DR PRINTS: PR00326; GTP1_OBG.
DR TIGRfams: TIGR00650; MG442; 2.
DR TIGRfams: TIGR00231; small_GTP; 2.
KW GTP-binding; Repeat; Complete proteome.
FT NP_BIND 10 17 GTP 1 (POTENTIAL).
FT NP_BIND 57 61 GTP 1 (POTENTIAL).
FT NP_BIND 119 122 GTP 1 (POTENTIAL).
FT NP_BIND 182 189 GTP 2 (POTENTIAL).
FT NP_BIND 229 233 GTP 2 (POTENTIAL).
FT NP_BIND 294 297 GTP 2 (POTENTIAL).
FT CONFLICT 186 187 VG -> CR (in Ref. 2).
SO SEQUENCE 436 AA; 48769 MW; ASCCT028FFBSA442 CRC64;

Query Match 76.8%; Score 1699; DB 1; Length 436;
Best Local Similarity 73.6%; Pred. No. 6.8e-97;
Matches 330; Conservative 57; Mismatches 58; Indels 0; gaps 0;

```

```

OY 181 IGRPNKSSLVNALLGEDRVYSNVACTTBDADITEXSYDGDYVLIOTACRKKKGY 240
DB 181 IGRPNKSSLVNALLGEDRVYSNVACTTBDADITEXSYDGDYVLIOTACRKKKGY 240
OY 241 ESTEKSVLRALKALERSNVVLVIDAEGITBODKRVAGYAHBOGKAVIYVNMKVYE 300
DB 241 ESTEKSVLRALKALERSNVVLVIDAEGITBODKRVAGYAHBOGKAVIYVNMKVYE 300
OY 301 KDKTKKKRDEVRKEFOFLDYAOIAFVSAREKTRLTLPPIYIEASENKKRVOSSSTLN 360
DB 301 KDKTKKKRDEVRKEFOFLDYAOIAFVSAREKTRLTLPPIYIEASENKKRVOSSSTLN 360
OY 301 KDESTKKEFEENRHOFLDYAPILFMSALTKRHTLPAILKASENSLRVOTNVN 360
DB 301 KDESTKKEFEENRHOFLDYAPILFMSALTKRHTLPAILKASENSLRVOTNVN 360
OY 361 EYVTDASINPPTDGRRLNVEYVAIEPPTFFVYVNDVLMHSTKRYLENDIRAA 420
DB 361 EYVTDASINPPTDGRRLNVEYVAIEPPTFFVYVNDVLMHSTKRYLENDIRAA 420
OY 361 DVIMDAVANNPPTTHNGSRKLYATVQVSVPSPFVVDPELHMFSTERYFLNNRI 420
DB 361 DVIMDAVANNPPTTHNGSRKLYATVQVSVPSPFVVDPELHMFSTERYFLNNRI 420
OY 421 FGEGGTPIHIIARRK 435
DB 421 FGEGGTPIHIIARRK 435
OY 421 FGEGGTPIHIIARRK 435
DB 421 FGEGGTPIHIIARRK 435

RESULT 5
ENCA_BACHD STANDARD: PRT: 437 AA.
ID ENCA_BACHD STANDARD: PRT: 437 AA.
AC Q9KCD4;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DE 28-FEB-2003 (Rel. 41, Last annotation update)
DE Probable GTP-binding protein enca.
OS ENCA OR BH1638.
OS Bacillus halodurans.
CC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=86655;
RN
RN SEQUENCE FROM N.A.
RP STRAIN-C-125 / JCM 9153;
RX MEDLINE-20512582; PubMed-11058132;
RA Takami H., Nakasone K., Takaki Y., Maeno G., Sasaki R., Masui N.,
RA Fuji F., Hirama C., Nakamura Y., Ogasawara N., Kuhara S.,
RA Horikoshi K.;
RT "Complete genome sequence of the alkaliphilic bacterium Bacillus
RT halodurans and genomic sequence comparison with Bacillus subtilis".
RL Nucleic Acids Res. 28:4317-4331(2000).
CC -1- SIMILARITY: BELONGS TO THE ERA/TIME FAMILY OF GTP-BINDING
CC PROTEINS. ENCA SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL: AP001512; BAB05357.1;
DR PIR: F83854; F83854.
DR HAMAP: MF_00195; -; 1.
DR InterPro: IPR005289; GTP-binding_dom.
DR InterPro: IPR006073; GTP1_OBG.
DR InterPro: IPR002917; MMR_HSR1.
DR InterPro: IPR005225; Small_GTP.
DR Pfam: PF01926; MMR_HSR1; 1.
DR PRINTS: PR00326; GTP1_OBG.
DR TIGRfams: TIGR00650; MG442; 2.
DR TIGRfams: TIGR00231; small_GTP; 2.
KW GTP-binding; Repeat; Complete proteome.
FT NP_BIND 10 17 GTP 1 (POTENTIAL).
FT NP_BIND 57 61 GTP 1 (POTENTIAL).
FT NP_BIND 119 122 GTP 1 (POTENTIAL).
FT NP_BIND 182 189 GTP 2 (POTENTIAL).
FT NP_BIND 229 233 GTP 2 (POTENTIAL).
FT NP_BIND 294 297 GTP 2 (POTENTIAL).
SO SEQUENCE 437 AA; 49024 MW; CIA4BD65A3ADAB CRC64;

```


OY	121	VDNMKREATDYDYSSLSGFEPEPIGSGSLGLGLDLAAVSHSGEEDPPYEDTILSL	180
OY	121	VDNMKREATDYDYSSLSGFEPEPIGSGSLGLGLDLAAVSHSGEEDPPYEDTILSL	180
DB	121	VDMPEKRNIDYDPSYSLGGLDPPYPSVSHVIGTGGVLDIAVENLPVEBAEEND-DILRS	179
OY	181	IGRPNGKSSSLVAILGEDRVYVSNVACTTDDAIDTEKS-YDDODVYLIDTGMKRGKV	239
DB	180	IGRPNGKSSSLVAILGEDRVYVSNVACTTDDAIDTEKS-YDDODVYLIDTGMKRGKV	239
OY	240	YESTEKYSVLRAKALAIERSNVVLVVIDAEQIIEQDRKVGAVAHEDGKAVYIVNKMDTV	299
DB	240	YESTEKYSVLRAKALAIERSNVVLVVIDAEQIIEQDRKVGAVAHEDGKAVYIVNKMDTV	299
OY	300	EXDSTKMPKREDEVRKEPEFLDYOAIARVSAKERTRLTPEPYINEASEMHRKRVOSSTL	359
DB	300	EXDSTKMPKREDEVRKEPEFLDYOAIARVSAKERTRLTPEPYINEASEMHRKRVOSSTL	359
OY	360	NVYVTDIAISNPTPTDKGRRLNVFYATOVAIEPPTEPVFNVDYELMHFKRYLEMOIRA	419
DB	360	NVYVTDIAISNPTPTDKGRRLNVFYATOVAIEPPTEPVFNVDYELMHFKRYLEMOIRA	419
OY	420	AFEGEGTPIHIIARKR 435	
DB	420	AFEGEGTPIHIIARKR 435	
RESULT 8			
ENGA_STRPN	STANDARD	PRT	436 AA.
AC	097PC9;		
DT	28-FEB-2003 (Rel. 41, Created)		
DT	28-FEB-2003 (Rel. 41, Last sequence update)		
DT	28-FEB-2003 (Rel. 41, Last annotation update)		
DE	Probable GTP-binding protein enga.		
GN	ENGA OR SERA OR SPI709 OR SPI553.		
OS	Streptococcus pneumoniae, and		
OS	Streptococcus pneumoniae (strain ATCC BAA-255 / 86).		
OC	Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;		
OC	Streptococcus.		
OX	NCBI_TaxID=1313, 171101;		
RM	11		
RP	SEQUENCE FROM N.A.		
RC	STRAIN-ATCC BAA-334 / TIGR4;		
RX	MEDLINE-21557209; PubMed-1163916;		
RA	Tettelein H., Nelson K.E., Paulsen I.T., Eisen J.A., Read T.D.,		
RA	Peterson S., Heidelberg J., Deboy R.T., Haft D.H., Dodson R.J.,		
RA	Durkin A.S., Gwinn M., Kolonay J.F., Nelson W.C., Peterson J.D.,		
RA	Umayam L.A., White O., Salzberg S.L., Lewis M.R., Redun D.,		
RA	Holtzapple E., Khouri H., Wolf A.M., Ullrich T.R., Hansen C.L.,		
RA	McDonald L.A., Feldblyum T.V., Angiuoli S., Dickinson T., Hickey E.K.,		
RA	Holt I.E., Loftus B.J., Yang F., Smith H.O., Venter J.C.,		
RA	Dougherty B.A., Morrison D.A., Hollingshead S.K., Fraser C.M.;		
RT	*Complete genome sequence of a virulent isolate of Streptococcus		
RT	pneumoniae.*		
RL	Science 293:498-506(2001).		
RM	121		
RP	SEQUENCE FROM N.A.		
RC	STRAIN-ATCC BAA-255 / R6;		
RX	MEDLINE-21429245; PubMed-11544234;		
RA	Hoskins J., Alborn M.E. Jr., Arnold J., Blaszcak L.C., Burgess S.,		
RA	Dehoff B.S., Estrem S.T., Filiz L., Fu D.-J., Fuller W., Gertinger C.,		
RA	Gillmore R., Glas J.S., Kholja H., Kraft A.R., Lagace R.E.,		
RA	Leclair D.J., Lee L.N., Lettowitz E.J., Lu J., Matsushima P.,		
RA	McMahon S.M., McHenry M., Mcleaster K., Mundy C.W., Nicas T.I.,		
RA	Norris F.H., O'Gara M., Peery R.B., Robertson G.T., Rockey P.,		
RA	Sun P.-M., Winkler M.E., Yang Y., Young-Bellido M., Zhao G.,		
RA	Zook C.A., Balz R.H., Jaskunas S.R., Rostek P.R. Jr., Skatrud P.L.,		
RA	Glass J.I.;		
RT	*Genome of the bacterium Streptococcus pneumoniae strain R6.*		
RL	J. Bacteriol. 183:5709-5717(2001)		
CC	-I- STRAINARTY: BELONGS TO THE ERA/TIME FAMILY OF GTP-BINDING		
CC	PROTEINS. ENGA SUBFAMILY.		
CC	-----		
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration		

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-slb.ch/announce/>
 CC or send an email to license@isb-slb.ch).

CC EMBL: AE007464: AK75787.1: -
 DR EMBL: AE008533: AK00357.1: -
 DR PIR: B95199; B95199.
 DR PIR: H98065; H98065.
 DR TIGR: SP1709; -
 DR HAMAP: MF_00195; - 1.
 DR InterPro: IPR005289; GTP-binding_dom.
 DR InterPro: IPR006073; GTP_OBG.
 DR InterPro: IPR002917; MMR_HSR1.
 DR InterPro: IPR005225; Small_GTP.
 DR Pfam: PF01926; MMR_HSR1. 1.
 DR PRINTS: PR00326; GTP_OBG.
 DR TIGR: TIGR00650; MG442; 2.
 DR TIGR: TIGR00231; small_GTP; 2.
 DR TIGR: TIGR00231; small_GTP; 2.
 KW GTP-binding; Repeat; Complete proteome.
 FT NP_BIND 10 17 GTP 1 (POTENTIAL).
 FT NP_BIND 57 61 GTP 1 (POTENTIAL).
 FT NP_BIND 119 122 GTP 1 (POTENTIAL).
 FT NP_BIND 181 188 GTP 2 (POTENTIAL).
 FT NP_BIND 229 233 GTP 2 (POTENTIAL).
 FT NP_BIND 294 297 GTP 2 (POTENTIAL).
 FT NP_BIND 394 397 GTP 2 (POTENTIAL).
 SO SEQUENCE 436 AA: 49082 MW: 80138900528585BDD CRC64:

Query Match 68.8% Score 1522; DB 1; Length 436;
 Best Local Similarity 66.2%; Pred. No. 4; 4e-86;

Matches 290; Conservative 73; Mismatches 69; Indels 6; Gaps 3;

OY 1 MTKPIVAVGRPNVSKSTIFNRIVGERSIVEDPTGVTDRIRYSSEGLTHDPIIIDGG 60
 DB 1 MALPTIAVGRPNVSKSTIFNRIVGERSIVEDPTGVTDRIRYATGEMHNSFSMIDGG 60
 OY 61 IEIGDAPFOTOIRAOAETIADEAVIIFPNVNRGSLGSDENVAOILYKSKRPVLA 120
 DB 61 IDVDAPFMDIQAQIAETIADEAVIIFPNVNRGSLGSDENVAOILYKSKRPVLA 120
 OY 121 VDNEMKRDVDFYSLGEGEPYPIGSHGGLGDLDAVVSHEGEEDEPDYEDTIRL 178
 DB 121 VDNEMKRDVDFYSLGEGEPYPIGSHGGLGDLDAVVSHEGEEDEPDYEDTIRL 178
 OY 121 VDNEMKRDVDFYSLGEGEPYPIGSHGGLGDLDAVVSHEGEEDEPDYEDTIRL 178
 DB 121 VDNEMKRDVDFYSLGEGEPYPIGSHGGLGDLDAVVSHEGEEDEPDYEDTIRL 178
 OY 179 SIIGRPVSKSLVNAIIGEDRIYVSNVAGTTRDAIDTFRYS-YDGDVYLIDTAGMRK 237
 DB 178 SIIGRPVSKSLVNAIIGEDRIYVSNVAGTTRDAIDTFRYS-YDGDVYLIDTAGMRK 237
 OY 238 KYESTERYSVLRALAIERSNVVLADEOGIIEODKRVAGVAHEDGKAVVIVVKN 297
 DB 238 KYESTERYSVLRALAIERSNVVLADEOGIIEODKRVAGVAHEDGKAVVIVVKN 297
 OY 238 KYESTERYSVLRALAIERSNVVLADEOGIIEODKRVAGVAHEDGKAVVIVVKN 297
 DB 238 KYESTERYSVLRALAIERSNVVLADEOGIIEODKRVAGVAHEDGKAVVIVVKN 297
 OY 298 TVEKDSKTKMKFEDEVAKKEFOFLDYAOIAFVSAKERTRLTLFPYINSESENHKKRV 357
 DB 298 TVEKDSKTKMKFEDEVAKKEFOFLDYAOIAFVSAKERTRLTLFPYINSESENHKKRV 357
 OY 298 TVEKDSKTKMKFEDEVAKKEFOFLDYAOIAFVSAKERTRLTLFPYINSESENHKKRV 357
 DB 298 TVEKDSKTKMKFEDEVAKKEFOFLDYAOIAFVSAKERTRLTLFPYINSESENHKKRV 357
 OY 358 TLNEVYDAISAMPTPDGKRLNVATOVATIEPTFVFNVDVLEHSEKRLLENOI 417
 DB 358 TLNEVYDAISAMPTPDGKRLNVATOVATIEPTFVFNVDVLEHSEKRLLENOI 417
 OY 358 TLNEVYDAISAMPTPDGKRLNVATOVATIEPTFVFNVDVLEHSEKRLLENOI 417
 DB 358 TLNEVYDAISAMPTPDGKRLNVATOVATIEPTFVFNVDVLEHSEKRLLENOI 417
 OY 418 RAAFGFEGTPIHIIARKR 435
 DB 418 RAAFGFEGTPIHIIARKR 435

RESULT 9
 ENCA_LACLA STANDARD; PRT: 436 AA.
 AC 09CHIG:
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Probable GTP-binding protein enga.
 GN ENGA OR YPHL OR IL0755.
 OS Lactococcus lactis (subsp. lactis) (Streptococcus lactis).
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae; Lactococcus.
 OX NCBI_TaxID=1360;
 RP SEQUENCE FROM N.A.
 RC STRAIN=IL1403;
 RA MEDLINE=2135186; PubMed=11337471;
 RA Bolotin A., Wincker P., Manger S., Jallou O., Malarme K.,
 RA Weissenbach J., Ehrlich S.D., Sorokin A.
 RA "The complete genome sequence of the lactic acid bacterium Lactococcus
 RT lactis ssp. lactis IL1403."
 RL Genome Res. 11:731-753(2001).
 CC -1- SIMILARITY: BELONGS TO THE ERA/TRME FAMILY OF GTP-BINDING
 CC PROTEINS. ENCA SUBFAMILY
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-slb.ch/announce/>
 CC or send an email to license@isb-slb.ch).

DR EMBL: AE006309: AK04853.1: -
 DR PIR: C86719; C86719.
 DR HAMAP: MF_00195; - 1.
 DR InterPro: IPR005289; GTP-binding_dom.
 DR InterPro: IPR006073; GTP_OBG.
 DR InterPro: IPR002917; MMR_HSR1.
 DR InterPro: IPR005225; Small_GTP.
 DR Pfam: PF01926; MMR_HSR1. 1.
 DR PRINTS: PR00326; GTP_OBG.
 DR TIGR: TIGR00650; MG442; 2.
 DR TIGR: TIGR00231; small_GTP; 2.
 KW GTP-binding; Repeat; Complete proteome.
 FT NP_BIND 10 17 GTP 1 (POTENTIAL).
 FT NP_BIND 57 61 GTP 1 (POTENTIAL).
 FT NP_BIND 119 122 GTP 1 (POTENTIAL).
 FT NP_BIND 181 188 GTP 2 (POTENTIAL).
 FT NP_BIND 229 233 GTP 2 (POTENTIAL).
 FT NP_BIND 294 297 GTP 2 (POTENTIAL).
 SO SEQUENCE 436 AA: 48866 MW: 1090081F966589D CRC64;

Query Match 68.7% Score 1520; DB 1; Length 436;
 Best Local Similarity 66.3%; Pred. No. 5; 9e-86;

Matches 289; Conservative 69; Mismatches 76; Indels 2; Gaps 2;

OY 1 MTKPIVAVGRPNVSKSTIFNRIVGERSIVEDPTGVTDRIRYSSEGLTHDPIIIDGG 60
 DB 1 MSIPVAVGRPNVSKSTIFNRIVGERSIVEDPTGVTDRIRYATGEMHNSFSMIDGG 60
 OY 61 IEIGDAPFOTOIRAOAETIADEAVIIFPNVNRGSLGSDENVAOILYKSKRPVLA 120
 DB 61 IELSDPEFTETIRAOAETIADEAVIIFPNVNRGSLGSDENVAOILYKSKRPVLA 120
 OY 121 VDNEMKRDVDFYSLGEGEPYPIGSHGGLGDLDAVVSHEGEEDEPDYEDTIRL 180
 DB 121 VDNEMKRDVDFYSLGEGEPYPIGSHGGLGDLDAVVSHEGEEDEPDYEDTIRL 180
 OY 121 VDNEMKRDVDFYSLGEGEPYPIGSHGGLGDLDAVVSHEGEEDEPDYEDTIRL 180
 DB 121 VDNEMKRDVDFYSLGEGEPYPIGSHGGLGDLDAVVSHEGEEDEPDYEDTIRL 180
 OY 181 IGRPNVSKSLVNAIIGEDRIYVSNVAGTTRDAIDTFRYS-SYGDVYLIDTAGMRK 239
 DB 180 IGRPNVSKSLVNAIIGEDRIYVSNVAGTTRDAIDTFRYS-SYGDVYLIDTAGMRK 239
 OY 181 IGRPNVSKSLVNAIIGEDRIYVSNVAGTTRDAIDTFRYS-SYGDVYLIDTAGMRK 239
 DB 180 IGRPNVSKSLVNAIIGEDRIYVSNVAGTTRDAIDTFRYS-SYGDVYLIDTAGMRK 239
 OY 240 YESTERYSVLRALAIERSNVVLADEOGIIEODKRVAGVAHEDGKAVVIVVKN 299
 DB 240 YESTERYSVLRALAIERSNVVLADEOGIIEODKRVAGVAHEDGKAVVIVVKN 299
 OY 240 YESTERYSVLRALAIERSNVVLADEOGIIEODKRVAGVAHEDGKAVVIVVKN 299
 DB 240 YESTERYSVLRALAIERSNVVLADEOGIIEODKRVAGVAHEDGKAVVIVVKN 299
 OY 300 EKDSKTKMKFEDEVAKKEFOFLDYAOIAFVSAKERTRLTLFPYINSESENHKKRV 359
 DB 300 EKDSKTKMKFEDEVAKKEFOFLDYAOIAFVSAKERTRLTLFPYINSESENHKKRV 359

```
QY 360 NEVYDAISMNPTDGRRLNRYATVOAIEPTFFVFNVDVLMHFSYKRLNIOIRA 419
DB 360 NDVIMDAVINPTDGRRLNRYATVOAIEPTFFVFNVDVLMHFSYKRLNIOIRA 419
QY 420 AFCECTPIHIIARRK 435
DB 420 AFCECTPVHLIARRK 435

RESULT 10
ENGA_CLOPE
ID ENGA_THETN STANDARD: PRT: 439 AA.
AC Q89JL;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DE 28-FEB-2003 (Rel. 41, Last annotation update)
DE Probable GTP-binding protein enga.
GN ENGA OR TTE1619.
OS Thermoaerobacter tengcongensis.
OC Bacteria; Firmicutes; Clostridia; Thermoaerobacteriales;
OC Thermoaerobacteriaceae; Thermoaerobacter.
OX NCBI_TaxID=119072;
RN 11
RP SEQUENCE FROM N.A.
RC STRAIN=MB4 / JCM 11007;
RX MEDLINE=21992816; PubMed=11997336;
RA Bao Q., Tian Y., Li W., Xu Z., Xuan Z., Hu S., Dong W., Yang J.,
RA Chen Y., Xue Y., Xu Y., Lai X., Huang L., Dong X., Ma Y., Ling L.,
RA Ten H., Chen R., Wang J., Yu J., Yang H.;
RT "A complete sequence of T. tengcongensis genome."
RL Genome Res. 13:689-700(2002)
CC -1- SIMILARITY: BELONGS TO THE ERA/TRME FAMILY OF GTP-BINDING
CC PROTEINS. ENGA SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
```

```
QY 119 NKVDNM-EMKTDVDFYSLGFGSPYPISGSHGLGDLDAVNSHFCEEDPYDEDTIR 177
DB 120 NKVDSFKEMKASFDYDFKGLGIPISASNGLGIGELDLDEINDEVE-YEEETIR 178
QY 178 LSIIGRPNGKSSLVNAILGEDRIVSNAGTTRADITREXSYGODVILDTGMRKKG 237
DB 179 IAVIGRPNGKSSLVNRIIGDEERVISDIPGTRADITPTKGRNIIIDTNGIRKS 238
QY 238 KYESTEKYSVLAKAIENSNVLYIDAEGIIIDOKRAGTAHEGKAAVIVYVKKMD 297
DB 239 RISEIERYSVLALAIERADICLMDATGTEGDTKAGYAFENGKGIILVKKMD 298
QY 298 IVEKDSKTKMKFEDEVKKEFOFLDYNOIAFYSAKERTRLTLPYINASESNHKKRVQSS 357
DB 299 IVEKDSNTYKEYEMIREKLAIFYSAFLIFISAKTGRIHVKLVLETVOKVMEVKKRTTG 358
QY 358 TLNFVYVDASINPTDGRRLNRYATVOAIEPTFFVFNVDVLMHFSYKRLNIOIRA 417
DB 359 LNNVNLNEMAKPPSPSSKGRPKITATVGTAKPFTVIFVNEBELHFSYVAFLENTI 418
QY 418 RAAFEGCTPIHIIARRK 435
DB 419 RONGFEGCTPIVISTKRK 436

RESULT 11
ENGA_CLOPE
ID ENGA_CLOPE STANDARD: PRT: 438 AA.
AC Q8YJKL;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DE 28-FEB-2003 (Rel. 41, Last annotation update)
DE Probable GTP-binding protein enga.
GN ENGA OR CPE1755.
OS Clostridium perfringens.
OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
OC Clostridium.
OX NCBI_TaxID=1502;
RN 11
RP SEQUENCE FROM N.A.
RC STRAIN=13 / Type A;
RX MEDLINE=21664373; PubMed=11792842;
RA Shimizu T., Onoani K., Hiraoka H., Ohshima K., Yamashita A.,
RA Shiba T., Ogasawara N., Hattori M., Kuhara S., Hayashi H.;
RT "Complete genome sequence of Clostridium perfringens, an anaerobic
RT flesh-eater."
RL Proc. Natl. Acad. Sci. U.S.A. 99:996-1001(2002).
CC -1- SIMILARITY: BELONGS TO THE ERA/TRME FAMILY OF GTP-BINDING
CC PROTEINS. ENGA SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
```



```

DR InterPro: IPR001806: Ras_trnsf_tnmg.
DR InterPro: IPR005225: Small_GTP.
DR Pfam: PF01976: MMR_HSR1. 1.
DR PRINTS: PRO0326: GTPIOBG.
DR PRINTS: PRO0449: RASTRANSFRMNG.
DR SMART: SM00382: AAA; 2.
DR TIGRFAMs: TIGR00650: MG442; 2.
DR TIGRFAMs: TIGR00231: small_GTP; 2.
KW GTP-binding; Repeat; Complete proteome.
FT NP-BIND 10 17 GTP 1 (POTENTIAL).
FT NP-BIND 57 61 GTP 1 (POTENTIAL).
FT NP-BIND 120 123 GTP 1 (POTENTIAL).
FT NP-BIND 183 190 GTP 2 (POTENTIAL).
FT NP-BIND 230 234 GTP 2 (POTENTIAL).
FT NP-BIND 295 298 GTP 2 (POTENTIAL).
SQ SEQUENCE 452 AA; 50825 MW; 7EF771EAF7DBCC76 CRC64;

Query Match 46.1% Score 1021; DB 1; Length 452;
Best Local Similarity 45.9% Pred. No. 1,9e-55;
Matches 200; Conservative 94; Mismatches 138; Indels 4; Gaps 4;

1 MFKPIYAIYGRNWKSTIFNRIVGERSVIEOPGVTNRIRYSSEGLTHDFENIIQCG 60
1 :|||||:|||||: 11 : :||| :|||: 1 :|||:
1 MSPLPIAIIIGRNWKSTFYRRLRLGNQQAIVHQDPIITRDRITRPFNFDRDPOVYVDG 60
1 :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
61 -EIG-DAPQTOIRQAQETAIDEADVIIENVRREGTQSDKEMVAQIILYKSKPPVILAVN 119
61 : :||| :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
61 LVFNDSSELPPIRQANLALAEAKAIFVVDGCGGPTASDEEIAQMLKQDSVPVILAVN 120
61 :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

```

```

Db      122 KCSPDQAIQAEEFMHJGLCEPYPMASIHGSGTCLDLALLETLPAPQEEP -BEDETKV 179

OY      179 STIGRPNKSKSLVAILIGEDBRVYVNAVGTBDADIDZESYSGODYVLDIADGRKKK 238
        :.:::|||||:||||: ||: |||: :.|||||: |||: |||||:|:|
Db      180 AIVGRPNVSKSSLALNAGEORALVSPISGTDADIDMYERNQKTYRLDTAGIRKKK 239
        :|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|

OY      239 VVESTKESYVLAALKAIERSNVLYVDAEGGIIIBQDKRAGYAHOGCAVYIVYKKDT 298
        :|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
Db      240 VDIAGFEFENAFIRADOVLYFLVDLDVGTEDBLKAGHILEDGAVVLYVINKDA 299
        :|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|

OY      299 VEKDSYTKMKRFEDEYRKEEFOFLDYAQAIYFVSAKERLRLTLPFYINEASENHKKRVSST 358
        :|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
Db      300 VEEEDSTYITEHEEQQLMARLYFMDNAEMIFVSAOTGRVOKIILDCVDIAOEHRRRRTYAV 359
        :|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|

OY      359 LNEVYVDAISNMPETDK -GRRLVVFATOVATIEPTFVFNVDVDELMFSYKRYLENOI 417
        :|||: :|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
Db      360 INVLELEANSMSPTTRGQKGIYCTQVSTOPALALFLVNDPFRFNDNRRRIENOF 419
        :|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|

OY      418 RAAFGEGTPIHITAR 433
        :|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
Db      420 RKOLGFGSPIRLFWR 435
        :|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|

Search completed: September 10, 2003, 00:33:14
Job time : 25 secs

```

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 10, 2003, 00:29:31; Search time 100 Seconds

(without alignments)
1125.109 Million cell updates/sec

Title: US-09-815-242-12600

Perfect score: 2213

Sequence: 1 MTKPIVAIVGRPNVGRKSTIF.....IRAPGFGPTPIHIAKRN 436

Scoring table: BIOSUM62

Gapop 10.0, Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 08

Maximum Match 1008

Listing first 45 summaries

Database: 1: SP archaea: 2: SP bacteria: 3: SP fungi: 4: SP human: 5: SP invertebrate: 6: SP mammal: 7: SP mnc: 8: SP organelle: 9: SP phage: 10: SP plant: 11: SP rodent: 12: SP virus: 13: SP vertebrate: 14: SP unclassified: 15: SP virus: 16: SP bacteriophage: 17: SP archaea: 18: SP archaea: 19: SP archaea: 20: SP archaea: 21: SP archaea: 22: SP archaea: 23: SP archaea: 24: SP archaea: 25: SP archaea: 26: SP archaea: 27: SP archaea: 28: SP archaea: 29: SP archaea: 30: SP archaea: 31: SP archaea: 32: SP archaea: 33: SP archaea: 34: SP archaea: 35: SP archaea: 36: SP archaea: 37: SP archaea: 38: SP archaea: 39: SP archaea: 40: SP archaea: 41: SP archaea: 42: SP archaea: 43: SP archaea: 44: SP archaea: 45: SP archaea

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2094	94.6	436	16	08CP62
2	1814	72.9	436	16	08EQ48
3	1523	68.8	436	16	08E379
4	1476	66.7	436	16	08DY73
5	1472	66.5	436	16	08RHV5
6	1472	66.5	436	16	08DS90
7	1342	60.6	436	16	08GES8
8	1340.5	60.6	436	16	08KH12
9	1028.5	46.5	449	16	08DK11
10	1022	46.2	444	16	08E3H5
11	920	41.6	503	16	08E7F5
12	898.5	40.6	496	16	08DF02
13	874	39.5	537	10	09LH58
14	874	39.5	659	10	09C7C0
15	859	38.8	487	16	08EC36
16	851.5	38.5	463	16	08G6A8

17	843.5	38.1	552	16	08E7K5	08E7K5 corynebacte
18	838	37.9	588	16	09A7R6	09A7R6 caulobacter
19	796.5	36.0	483	16	08G2E8	08G2E8 brucella su
20	756.5	34.2	874	5	08ISN5	08ISN5 plasmodium
21	736	33.3	496	10	09A7R4	09A7R4 guillardi
22	678	30.6	489	16	08E6K1	08E6K1 leptospira
23	661	29.9	188	11	09E2I2	09E2I2 lactospora
24	559.5	25.3	383	2	P94645	P94645 campylobact
25	524	23.7	642	10	09A4X0	09A4X0 oryza sativ
26	514.5	23.2	514	10	09ELED	09ELED arabidopsi
27	512	23.1	347	2	08VN97	08VN97 helicobacte
28	512	23.1	347	2	08VN36	08VN36 helicobacte
29	511	23.1	347	2	08VN38	08VN38 helicobacte
30	455	20.6	281	2	09LA73	09LA73 thiolobact
31	435	19.7	456	16	08D1Y0	08D1Y0 virogenes
32	433.5	19.6	208	2	P72518	P72518 streptococ
33	320.5	14.5	190	2	09AE38	09AE38 campylobact
34	320.5	14.5	190	2	09AQ16	09AQ16 campylobact
35	319.5	14.4	190	2	09AE36	09AE36 campylobact
36	318.5	14.4	190	2	09AE39	09AE39 campylobact
37	316.5	14.3	190	2	09AE37	09AE37 campylobact
38	266	13.0	479	16	08DPZ8	08DPZ8 streptococ
39	258	11.7	300	16	08EPY0	08EPY0 oceanobact
40	243.5	11.0	455	16	08DPT8	08DPT8 streptococ
41	241.5	10.9	170	2	09S3C9	09S3C9 helicobacte
42	241.5	10.9	170	2	09S3C8	09S3C8 helicobacte
43	241.5	10.9	170	2	09S3C7	09S3C7 helicobacte
44	241.5	10.9	170	2	08VNA0	08VNA0 helicobacte
45	241.5	10.9	170	2	09R3H9	09R3H9 helicobacte

ALIGNMENTS

RESULT 1

ID	08CP62	PRELIMINARY:	PRT:	436 AA.
AC	08CP62:	01-MAR-2003 (TREMBLrel. 23, Created)		
DT	01-MAR-2003 (TREMBLrel. 23, Last sequence update)			
DT	01-MAR-2003 (TREMBLrel. 23, Last sequence update)			
DE	GN binding protein.			
GN	SE1163.			
OS	Staphylococcus epidermidis.			
OC	Bacteria: Firmicutes: Bacillales: Staphylococcus.			
OX	NCBI_TaxID=1282;			
RN	(1)			
RP	SEQUENCE FROM N.A.			
RC	STRAIN-ATCC 12228;			
RA	Zhang Y., Ren S., Li H., Fu G., Lu L., Lu G., Jia J., Yu Y., Qin Z.,			
RA	Chen Z., Men Y.;			
RL	Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.			
DR	EMBL: AE016747; AAC04760.1; -			
KM	Complete proteome.			
SO	SEQUENCE 436 AA; 49024 MM; 77C74B95641D3F22 CRC64;			

Query Match 94.64; Score 2094; DB 16; Length 436;

Best Local Similarity 93.18; Pred. No. 2.9e-124;

Matches 406; Conservative 18; Mismatches 12; Indels 0; Gaps 0;

QY	1	MTKPIVAIVGRPNVGRKSTIFNRIVGERVSIVEDPCVTDRIYSGGEMLTDFNIDTGC	60
DB	1	MTKPIVAIVGRPNVGRKSTIFNRIVGERVSIVEDPCVTDRIYSGGEMLTDFNIDTGC	60
QY	61	ITIGAPAPOTIRAOEIAIDEAVITIFNNVREGLTQSDENVAOILYKSKPPVLAANK	120
DB	61	ITIGAPAPOTIRAOEIAIDEAVITIFNNVREGLTQSDENVAOILYKSKPPVLAANK	120
QY	61	ITIGAPAPOTIRAOEIAIDEAVITIFNNVREGLTQSDENVAOILYKSKPPVLAANK	120
DB	61	ITIGAPAPOTIRAOEIAIDEAVITIFNNVREGLTQSDENVAOILYKSKPPVLAANK	120
QY	121	VNNMERMVDVDFSLGCEPPISSHGGLGDLDAVVSHPGEEEDPDYDOTTIRLSI	180
DB	121	VNNMERMVDVDFSLGCEPPISSHGGLGDLDAVVSHPGEEEDPDYDOTTIRLSI	180
QY	121	VNNMERMVDVDFSLGCEPPISSHGGLGDLDAVVSHPGEEEDPDYDOTTIRLSI	180
DB	121	VNNMERMVDVDFSLGCEPPISSHGGLGDLDAVVSHPGEEEDPDYDOTTIRLSI	180
QY	181	ICRPVNGKSLVNNILGEDRVIVSNVAGTTRDAIDREYSYDGDVVLIDTQGMRRKKVY	240

DB 181 IGRPNVGSLLVNAALGGERIVISNAGTTBDALDITETSTGDDVYLIDTAGKRRKKGY 240
OY 241 ESTERYSVLRALKAIERSNVVLYIDAEGCIIEODKRVAGYAHBOGKAVIYVNMKDTVE 300
DB 241 ESTERYSVLRALKAIERSNVVLYIDAEGCIIEODKRVAGYAHBOGKAVIYVNMKDTVE 300
OY 301 KDSKTKMKKFEDEVRKEFOFLDYAOIAFVSAKERTRLTLPFYINEASENHKKRVOSTLN 360
DB 301 KDSKTKMKKFEDEVRKEFOFLDYAOIAFVSAKERTRLTLPFYINEASENHKKRVOSTLN 360
OY 361 EYVTDALISNMPPTDCKGRNLNVFATOVATIEPTTFVFNVDVLMHFSKRYLENDIRAA 420
DB 361 EYVTDALISNMPPTDCKGRNLNVFATOVATIEPTTFVFNVDVLMHFSKRYLENDIRAA 420
OY 421 FGEFGTPIHIARKN 436
DB 421 FGEFGTPIHIARKN 436

RESULT 2

DB 08E0A8 PRELIMINARY: PRT: 436 AA.
AC 08E0A8: 01-MAR-2003 (TREMBlrel. 23, Created)
DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)
DE 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
GN O81797.
OS Oceanobacillus theysensis.
OC Bacteria: Firmicutes: Bacillales: Oceanobacillus.
OX NCBI_Taxid=182710.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-HTE831 / DSM 14371 / JCM 11309;
RX MEDLINE-2220767; PubMed-12235376;
RA Takami H., Takaki Y., Uchiyama I.;
RT "Genome sequence of Oceanobacillus theysensis isolated from the Iheya
RT Ridge and its unexpected adaptive capabilities to extreme
RT environments.";
RL Nucleic Acids Res. 30:3927-3935 (2002).
DR EMBL: AP004599; BAC13753.1; -
KW Complete proteome.
SO SEQUENCE 436 AA; 48998 MW; 18F52E40D9B0C9C2 CRC64;

Query Match 72.9%; Score 1614; DB 16; Length 436;
Best Local Similarity 69.4%; Pred. No. 5.8e-94;
Matches 302; Conservative 66; Mismatches 67; Indels 0; Gaps 0;

OY 1 MTKPIVATIGRPNVGSSTIFNRIIGERVSIVEDTPGVTRORIYSSGEMLTTHDNIITDGC 60
DB 1 MTKSVVATIGRPNVGSSTIFNRIIGERVSIVEDTPGVTRORIYSSGEMLTTHDNIITDGC 60
OY 61 IETGAPFOTOIRAOAEIAIDEADVIIFMNVNREGCTOSDEMVAOILYKSKRPVLAANK 120
DB 61 IETGAPFOTOIRAOAEIAIDEADVIIFMNVNREGCTOSDEMVAOILYKSKRPVLAANK 120
OY 121 VDNMEKRTDVFYSLGFEPEYPISSGSHGLGLDLDAVVAHSGEEDPEYDRTIRLSI 180
DB 121 VDNMEKRTDVFYSLGFEPEYPISSGSHGLGLDLDAVVAHSGEEDPEYDRTIRLSI 180
OY 181 IGRPNVGSLLVNAALGGERIVISNAGTTBDALDITETSTGDDVYLIDTAGKRRKKGY 240
DB 181 IGRPNVGSLLVNAALGGERIVISNAGTTBDALDITETSTGDDVYLIDTAGKRRKKGY 240
OY 241 ESTERYSVLRALKAIERSNVVLYIDAEGCIIEODKRVAGYAHBOGKAVIYVNMKDTVE 300
DB 241 ESTERYSVLRALKAIERSNVVLYIDAEGCIIEODKRVAGYAHBOGKAVIYVNMKDTVE 300
OY 301 KDSKTKMKKFEDEVRKEFOFLDYAOIAFVSAKERTRLTLPFYINEASENHKKRVOSTLN 360
DB 301 KDSKTKMKKFEDEVRKEFOFLDYAOIAFVSAKERTRLTLPFYINEASENHKKRVOSTLN 360

OY 361 EYVTDALISNMPPTDCKGRNLNVFATOVATIEPTTFVFNVDVLMHFSKRYLENDIRAA 420
DB 361 EYVTDALISNMPPTDCKGRNLNVFATOVATIEPTTFVFNVDVLMHFSKRYLENDIRAA 420
OY 421 FGEFGTPIHIARKN 435
DB 421 FGEFGTPIHIARKN 435

RESULT 3

DB 08E379 PRELIMINARY: PRT: 436 AA.
AC 08E379: 01-MAR-2003 (TREMBlrel. 23, Created)
DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)
DE 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
GN G851667.
OS Streptococcus agalactiae (serotype III).
OC Bacteria: Firmicutes: Lactobacillales: Streptococcaceae;
OX NCBI_Taxid=216495;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-NEM316 / Serotype III;
RX MEDLINE-2242508; PubMed-12354221;
RA Glaser P., Rusnlok C., Buchrieser C., Chevalier F., Frangeul L.,
RA Masdek T., Zouine M., Couve E., Lalloul L., Poyart C., Tillet-Cuot P.,
RA Kunst F.;
RT "Genome sequence of Streptococcus agalactiae, a pathogen causing
RT invasive neonatal disease.";
RL MOL. Microbiol. 45:1499-1513 (2002).
DR EMBL: AL766852; CAD47326.1; -
KW Hypothetical protein; Complete proteome.
SO SEQUENCE 436 AA; 48880 MW; E7E33B40F3F610CRC64;

Query Match 68.8%; Score 1523; DB 16; Length 436;
Best Local Similarity 66.5%; Pred. No. 3.2e-88;
Matches 290; Conservative 70; Mismatches 74; Indels 2; Gaps 2;

OY 1 MTKPIVATIGRPNVGSSTIFNRIIGERVSIVEDTPGVTRORIYSSGEMLTTHDNIITDGC 60
DB 1 MTKSVVATIGRPNVGSSTIFNRIIGERVSIVEDTPGVTRORIYSSGEMLTTHDNIITDGC 60
OY 61 IETGAPFOTOIRAOAEIAIDEADVIIFMNVNREGCTOSDEMVAOILYKSKRPVLAANK 120
DB 61 IETGAPFOTOIRAOAEIAIDEADVIIFMNVNREGCTOSDEMVAOILYKSKRPVLAANK 120
OY 121 VDNMEKRTDVFYSLGFEPEYPISSGSHGLGLDLDAVVAHSGEEDPEYDRTIRLSI 180
DB 121 VDNMEKRTDVFYSLGFEPEYPISSGSHGLGLDLDAVVAHSGEEDPEYDRTIRLSI 180
OY 181 IGRPNVGSLLVNAALGGERIVISNAGTTBDALDITETSTGDDVYLIDTAGKRRKKGY 240
DB 181 IGRPNVGSLLVNAALGGERIVISNAGTTBDALDITETSTGDDVYLIDTAGKRRKKGY 240
OY 241 ESTERYSVLRALKAIERSNVVLYIDAEGCIIEODKRVAGYAHBOGKAVIYVNMKDTVE 300
DB 241 ESTERYSVLRALKAIERSNVVLYIDAEGCIIEODKRVAGYAHBOGKAVIYVNMKDTVE 300
OY 301 KDSKTKMKKFEDEVRKEFOFLDYAOIAFVSAKERTRLTLPFYINEASENHKKRVOSTLN 360
DB 301 KDSKTKMKKFEDEVRKEFOFLDYAOIAFVSAKERTRLTLPFYINEASENHKKRVOSTLN 360
OY 361 EYVTDALISNMPPTDCKGRNLNVFATOVATIEPTTFVFNVDVLMHFSKRYLENDIRAA 420
DB 361 EYVTDALISNMPPTDCKGRNLNVFATOVATIEPTTFVFNVDVLMHFSKRYLENDIRAA 420
OY 420 AFGFECTPIHIARKN 435
DB 420 AFGFECTPIHIARKN 435

RESULT 4	08BDY73	PRELIMINARY:	PRT:	436 AA.
ID	08BDY73			
AC	08BDY73:			
DT	01-MAR-2003 (TREMBLrel. 23, Created)			
DT	01-MAR-2003 (TREMBLrel. 23, Last sequence update)			
DT	01-MAR-2003 (TREMBLrel. 23, Last annotation update)			
DE	Phosphoglycerate dehydrogenase-related protein.			
GN	SGA1620.			
OS	<i>Streptococcus agalactiae</i> (serotype V).			
OC	Bacteria: Firmicutes: Lactobacillales: Streptococcaceae:			
OC	<i>Streptococcus</i> .			
OC	<i>Streptococcus</i> .			
OX	NCBI_Taxid-216466;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	SPRAIN-2603 V/R / Serotype V:			
RX	MEDLINE-22223288; PubMed-12200547;			
RA	Tetzelin H., Maignien V., Cieslewicz M.J., Eisen J.A., Peterson S.,			
RA	Wessels L.R., Paulsen I.T., Nelson K.E., Margalit I., Read T.D.,			
RA	Maddif L.C., Wolf A.M., Beanan M.J., Brinkac L.M., Daugherty S.C.,			
RA	Deboy R.T., Durkin A.S., Kolony J.F., Madupu R., Lewis M.R.,			
RA	Radune D., Fedorova N.B., Scanlan D., Khouri H., Mulligan S.,			
RA	Carly H.A., Cline R.T., Van Aken S.E., Gill J., Scarselli M., Mora M.,			
RA	Iacobini E.T., Brettoni C., Gall G., Mariani M., Vegni F., Malone D.,			
RA	Ramado D., Raspoll R., Telford J.L., Kasper D.L., Grandi G.,			
RA	Fraser C.M.;			
RT	"Complete genome sequence and comparative genomic analysis of an			
RT	emerging human pathogen, serotype V <i>Streptococcus agalactiae</i> ."			
RL	Proc. Natl. Acad. Sci. U.S.A. 99:12391-12396(2002).			
DR	EMBL: AA014265; AA00484.1; -.			
DR	TrEM: SGA1620; -.			
DR	TrEM: SGA1620; -.			
QO	Complete proteome.			
QO	SEQUENCE 436 AA; 48981 MW; E2C064B3F04B8644 CRC64;			

Query Match	68.8%	Score 1523	DB 16	Length 436
Best Local Similarity	65.3%	Pred. No. 3,268		
Matches 289	Conservative 72	Mismatches 73	Indels 2	Gaps 2
QY	1	MTKPIYAIYGRPNVCKSTIFNRIYGERVSYEDTPTGYTRDRIYSSGGMLTGHNIDTGG	60	
Db	1	MTLPYAIYGRPNVCKSTIFNRIYGERVSYEDTPTGYTRDRIYSSGGMLTGHNIDTGG	60	
QY	61	IEIGDAPFQOTIRQAQAEIAIDEDADYIIFNVVREGILQSDDEMYAQILYKSKKPFVYLAIVK	120	
Db	61	IDDDVDAPEMQKHQADIAIMTEADYIVPVGSGKGVTDADDEYYSRLIYKKNKPVLAIVK	120	
QY	121	VDMEMKRTVVOFYSLGFCPEPTPIGSGHGLGGLDLDAVNSHSGEEDPDYEDDTRLSI	180	
Db	121	VDMPEKRNDIYFYSGLGDDPPLSVGICDTGILAIYENLPEVEENE-NDIIRFSL	179	
QY	181	IGRPVNGKSSLNALIGEDRIVYSNVACTTIDAIDTEY-XYGDGVYLLIPTAQRKKGV	239	
Db	180	IGRPVNGKSSLNALIGEDRIVYASPVACTTIDAIDTFNVSOGGEYTIMDTAGMRKSGV	239	
QY	240	YESETEYSYLRALKAIERSNVVLVYIDAEQGIITQDKRVAGVAGHEQKAVIYVNVKMDY	299	
Db	240	YESETEYSYMRKRAIDRSQDVLYATINAEQGIITQDKRVAGVAGHEQKGIITYVNMMDY	299	
QY	300	EXDKSTYMKFEDVEKKEQFLDYAOIAFVSAKERTRLRTLPYINBAENHKRVSSTL	359	
Db	300	EXKNHIVSOMGEIDIRNFQFLSYAIIIFVSAEQRLKLPDMKIRISQSNKRISAVL	359	
QY	360	NEVYTDAIMNPTPTDKGRLNVFAATOVAALEPTFVFPVNDVLEMFYSKRYLEQIRA	419	
Db	360	NOVYNDAIMNPTPTDKGRLKIFATOVAAKPTFPVFPVNEEELHFSRLRLNQIRE	419	
QY	420	AFGEFECTPHIATARR	435	
Db	420	AFVECTPTNLATARR	435	

```

09RHV5
ID O9RHV5 PRELIMINARY; PRT; 436 AA.
AC O9RHV5;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, last annotation update)
DE Phosphoglycerate dehydrogenase.
GN PGDA.
OS Streptococcus mutans.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1309;
RN 11
RP SEQUENCE FROM N.A.
RC STRAIN-MT8148;
RA Kawabata S., Terao Y., Hamada S.;
RT "Molecular cloning, sequence and characterization of a novel
RL streptococcal phosphoglycerate dehydrogenase gene.";
RL Oral Microbiol. Immunol. 15:58-62(2001).
DR EMBL; AB016077; BA88823.1; -.
DR InterPro: IPR005289; GTP-binding_dom.
DR InterPro: IPR006073; GTP1_OBG.
DR InterPro: IPR002917; MMR_HSR1.
DR InterPro: IPR005225; Small_GTP.
DR Pfam: PF01926; MMR_HSR1; 1.
DR PRINTS: PR00326; GTP1OBG.
DR TIGRFAMs: TIGR00650; MG442; 2.
DR TIGRFAMs: TIGR00231; small_GTP; 2.
DR SEQUENCE 436 AA; 48601 MW; EC2C8650D5D090 CRC04;
Query Match 66.7%; Score 1476; DB 2; Length 436;
Best Local Similarity 62.8%; Pseq. No. 3e-85; 77; Indels 2; Gaps
Matches 274; Conservative

```

Qy	1	MTEPIATVIGRRNKKSTLPIRNIGCEPVSYEDPRGVTDRRIYSCEGMLLHONPIITDGC	60
Db	1 <td>MALEPVALIGRRNKGSALEPNRIAGSERISTIVEDGVTDRITAKMLEROSTIIITDGC<td>60</td></td>	MALEPVALIGRRNKGSALEPNRIAGSERISTIVEDGVTDRITAKMLEROSTIIITDGC <td>60</td>	60
Qy	61	IEIGDAPFOTOIRAGAEIAIDEADVIIFPVNVRREGLTOSDEYAOILYKSKRPVLAUNK <td>120</td>	120
Db	61	IDOVDPAREMOQIKHOADIAMTEADYIVHVSASKEGIDAOAEYAKILLYRHKRVILAUNK <td>120</td>	120
Qy	121	VDMHEARTOVUYEYSIGFCEGPPIGSGIGLGGDLADAVYSHGEEDEEDPYDETRLSI <td>180</td>	180
Db	121	VDMPEKRSATVYEFYALGIGLGDPRPYPSAIGITGCGVDATYDNLPALEQEE-SSDIITKSTL <td>179</td>	179
Qy	181	IGRPNVGKSSLVNAIILEDGRVIVSVNAGTTRDAIOTREYS-YDGDYVLIDTAGMRKKGV <td>239</td>	239
Db	180	IGRPNGKSSLLINALIGEDRVIVASVPACTDADIDTFTDECEQEFMLIDTAGMRKSGV <td>239</td>	239
Qy	240	YESTEKYSVLRLAKAIERSNVVLYVIOAEQGLIEODKRVGVANRQCKANVYVNNKMDVY <td>299</td>	299
Db	240	YETETESYVMRMRIRIDSDIYLVLANNEEGRIDKRIKRGPHHEAGKGIYVYNNKMDVY <td>299</td>	299
Qy	300	ENDSKTKMKFEDEVKREFQPLDYAOIAFVSAKERTRLTLPYINAESENHKKRVQOSTL <td>359</td>	359
Db	300	KKNRNRVAQMEEDIDNFOQYIPYAPVYFASVNTQKRLHKLPDIKQVOSOMTRIPSSVL <td>359</td>	359
Qy	360	NEVVTATASINPTPTDKDGRRLNVFAATOVALEPTFVFNVDVLELHFSKRYLENOIQA <td>419</td>	419
Db	360	NDVYMDAIVAINPTPTDKGRKLKIFIAATOVSAKPTFVIFVNEELMHFSYLAFLENOITQ <td>419</td>	419
Qy	420	AFEGEGCTPIHIIATARR 435	
Db	420	AFVEECTPIRLIARRK 435	

RESULT 6

Q8DS90 ID Q8DS90 PRELIMINARY: PRT: 436 AA.

AC Q8DS90;

DT 01-MAR-2003 (TtEMblrel. 23, Created)

DT 01-MAR-2003 (TtEMblrel. 23, Last sequence update)

1

2

3

4

5

6

7

8

9

10

11

12

13

14

15

16

17

18

19

20

21

22

23

24

25

26

27

28

29

30

31

32

33

34

35

36

37

38

39

40

41

42

43

44

45

46

47

48

49

50

51

52

53

54

55

56

57

58

59

60

61

62

63

64

65

66

67

68

69

70

71

72

73

74

75

76

77

78

79

80

81

82

83

84

85

86

87

88

89

90

91

92

93

94

95

96

97

98

99

100

DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)
 DE Phosphoglycerate dehydrogenase.
 GN PGDA OR SMU.1920.
 OS Streptococcus mutans.
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 CC Streptococcus.
 NCBI_TaxID=1309;
 RX NCBI [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=UA159 / ATCC 700610 / Serotype C;
 RA Aylde D., McShan W.M., McLaughlin R.E., Savic G., Chang J.,
 RA Carson M.B., Primeaux C., Tian R., Kenton S., Jia H., Lin S., Qian Y.,
 RA Li S., Zhu H., Majar F., Lai H., White J., Roe B.A., Ferretti J.J.;
 RT "Genome sequence of Streptococcus mutans UA159, a cariogenic dental
 RT pathogen."
 RL Proc. Natl. Acad. Sci. U.S.A. 99:14434-14439(2002).
 DR EMBL: AE015016; AAN59531.1; -
 KM Complete proteome.
 SQ SEQUENCE 436 AA; 48585 MW; 2408428A91C2A097 CRC64;

Query Match 66.5%; Score 1472; DB 16; Length 436;
 Best Local Similarity 62.6%; Pred. No. 5.3e-85;
 Matches 273; Conservative 84; Mismatches 77; Indels 2; Gaps 2;

OY 1 MTKPIVAIYGRPNVSKSTIFNRIGERSYIEDTPGVTRDRIRYSSGEMLTNHDNIIDTGG 60
 DB 1 MALPTVAIYGRPNVSKSLAFNRINGEISTIEDVEGVTRITTKAMLNROPSIIDTGG 60
 OY 61 IEIGAPFOTOIRAOAEIAIDEADVIIFMNVNREGLTOSDEMAOILYKSKRPVLAIVK 120
 DB 61 IDVDVAPFEOIKHDAIDAMTEADVIIVVSAKSGITDADEYAKILYRHKRVILAIVK 120
 OY 121 VDNMEMRTDYDFYSLGEGEPYISGSHGLGDLDAVYSHGEEEDPYEDTIRLSI 180
 DB 121 VDNPEMSAIYDFYALGLGDPYVPSAHGIGTGDVLAIVDNLPTEAOEE-SSDIKFSL 179
 OY 181 ICRPNVGSLLVNAITLGEDRVIVSNVAGTTRDAIDTETS-YGODVYLIDTQMRKKRV 239
 DB 180 ICRPNVGSLLVNAITLGEDRVIVSNVAGTTRDAIDTETS-YGODVYLIDTQMRKKRV 239
 OY 240 YESTEKSYVALAKAIEISNVLVVIADEGIIIDOKRVAAGVIAHEGKAAVYVYVKNKDV 299
 DB 240 YESTEKSYVALAKAIEISNVLVVIADEGIIIDOKRVAAGVIAHEGKAAVYVYVKNKDV 299
 OY 300 EKDSKTKKFEDEVERKEFOFLDVAQIAVSAKERTRLTLFPIYINASENKKRVOSTL 359
 DB 300 KDNRTVAOWEADIDRNFQIYIPAPIVFSVAVTOKRLKLPDIYKOVSOQNRIPSAVL 359
 OY 360 NEVMTDAISMNPTPTDKGRRLNVYATOVAIIEPTFVFNVDVLELHFSKRYKLENOIRA 419
 DB 360 NOVVDVAIVNPTPTDKGRRLNVYATOVAIIEPTFVFNVDVLELHFSKRYKLENOIRA 419
 OY 420 AFGEGCTPIHITARRK 435
 DB 420 AFGEGCTPIHITARRK 435

RESULT 7

08G58 PRELIMINARY: PRT: 445 AA.
 AC 08G58:
 DT 01-MAR-2003 (Tremblrel. 23, Created)
 DT 01-MAR-2003 (Tremblrel. 23, Last sequence update)
 DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)
 DE GTP-binding protein (Fragment).
 OS Hellobacillus mobilis.
 OC Bacteria; Firmicutes; Clostridia; Clostridiales; Hellobacteriaceae;
 CC Hellobacillus.
 NCBI_TaxID=28064;
 RX NCBI [1]
 RP SEQUENCE FROM N.A.
 DR MEDLINE=22337798; PubMed=12446909;

RA Raymond J., Zhaxybayeva O., Gogarten J.P., Gerdes S.Y.,
 RA Blankenship R.E.;
 RT "Whole-genome analysis of photosynthetic prokaryotes."
 RL Science 298:1616-1620(2002).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Liollos K.G., Chu L., Ostrovskaya O., Mendybaeva N., Koukharenko V.,
 RA Gerdes S., Kyrides N., Overbeck R.;
 RL Submitted (Aug-2002) to the EMBL/Genbank/DBJ databases.
 DR EMBL: AY142760; AAN87364.1; -
 FT NON-TER 445 445
 SQ SEQUENCE 445 AA; 49885 MW; 1D6A692E387AD61C CRC64;

Query Match 60.6%; Score 1342; DB 2; Length 445;
 Best Local Similarity 56.8%; Pred. No. 8.8e-77;
 Matches 250; Conservative 83; Mismatches 99; Indels 8; Gaps 3;

OY 1 MTKPIVAIYGRPNVSKSTIFNRIGERSYIEDTPGVTRDRIRYSSGEMLTNHDNIIDTGG 60
 DB 4 MAPPIVAVGRPNVSKSTIFNRIGERSYIEDTPGVTRDRIRYSSGEMLTNHDNIIDTGG 63
 OY 61 IEIG--DAPFOGIRAOAEIAIDEADVIIFMNVNREGLTOSDEMAOILYKSKRPVLAIV 118
 DB 64 LEFGAGGNFSEVIYKQAEIAIDEADVIIFMNVNREGLTOSDEMAOILYKSKRPVLAIV 123
 OY 119 NKVDNMEMRTDYDFYSLGEGEPYISGSHGLGDLDAVYSHGEEEDPYEDTIRLSI 175
 DB 124 NKIEDSQDKYIEFPALGIDPPIISATGHMNGDLDVAIVELAPREGDEDDP---DT 180
 OY 176 IRLSTIIGRPNVSKSLVNAITLGEDRVIVSNVAGTTRDAIDTETS-YGODVYLIDTQMR 235
 DB 181 IRLSTIIGRPNVSKSLVNAITLGEDRVIVSNVAGTTRDAIDTETS-YGODVYLIDTQMR 240
 OY 236 KGVVYESTEKSYVALAKAIEISNVLVVIADEGIIIDOKRVAAGVIAHEGKAAVYVYVKN 295
 DB 241 KGVVYESTEKSYVALAKAIEISNVLVVIADEGIIIDOKRVAAGVIAHEGKAAVYVYVKN 300
 OY 296 MVDVDSKTKKFEDEVERKEFOFLDVAQIAVSAKERTRLTLFPIYINASENKKRVOSTL 355
 DB 301 MVDVDSKTKKFEDEVERKEFOFLDVAQIAVSAKERTRLTLFPIYINASENKKRVOSTL 360
 OY 356 SSTLNEVMTDAISMNPTPTDKGRRLNVYATOVAIIEPTFVFNVDVLELHFSKRYKLE 415
 DB 361 SSTLNEVMTDAISMNPTPTDKGRRLNVYATOVAIIEPTFVFNVDVLELHFSKRYKLE 420
 OY 416 OIRAFEGCTPIHITARRK 435
 DB 421 RFRFTGFGCTPIHITARRK 440

RESULT 8

08KH12 PRELIMINARY: PRT: 435 AA.
 AC 08KH12:
 DT 01-OCT-2002 (Tremblrel. 22, Created)
 DT 01-OCT-2002 (Tremblrel. 22, Last sequence update)
 DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)
 DE Conserved hypothetical GTP-binding protein.
 OS Lactobacillus delbrueckii (subsp. bulgaricus).
 OC Bacteria; Firmicutes; Lactobacillales; Lactobacillaceae;
 CC Lactobacillus.
 NCBI_TaxID=1585;
 RX NCBI [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 11842;
 RA Serrif P., Deruyin R., Ehrlich S.D., Maguin E.;
 RT "Lactobacillus delbrueckii spp. bulgaricus h1da region."
 RL Submitted (Apr-2002) to the EMBL/Genbank/DBJ databases.
 DR EMBL: AY094626; AAM22484.1; -
 DR InterPro: IPR000795; EF-GTPbind.
 DR InterPro: IPR005289; GTP-binding_dom.
 DR InterPro: IPR006073; GTP_OBG.
 DR InterPro: IPR002917; MMR_HSR1.

QY 125 EMRT-DVDFYSLGSGEPYPIGSHGGLGLDLDAVSHSGEEDDPIDETIRLSITGR 183
DB 124 NKOSIDHEHLEHFGKPFISANHGIGGLDLIVD---TKLEKEEDTFFECITGR 180
QY 184 PNAGSSLVNALIGEDRVIVSNVACTRDAIDTETVSYDGODYVLIDTAGMKKKQYEST 243
DB 181 PNAGSSLVNACILNEDRMITSNIANTIRDAIDSNEKNDLTYIIDTAGIRKKKIDENV 240
QY 244 EKYSVLRAKAEISRNVLVVIDAEOGIIEDOKRAGVAGHOGKAVIYVNMKDDY-EKD 302
DB 241 DKAVILRVGOSISRSRLIYVLDGSEFNEGDEVIAGLAHKANIPSLIYVNMKDIYKED 300
QY 303 SKTKMKFFEDVREKPEFLDYAOIAFVSAKERTRLTLEPTYNESNMKKRQVOSTLNEY 362
DB 301 EKTMKMFITIRIKFELWTPIVFLSLAENKRIATLFEKISIREMLNLMKFSKITLTDL 360
QY 363 VTDAISMNPTPTDKGRNLNMFYATOVAIEPTFVYVNDVLMHFSYKRYLENOIRAFG 422
DB 361 VFKLOMLNSPPLNRRGRKINHVTOVGOIPTFVLCNPEYLFHFSYARYLENEIRKSLG 420
QY 423 FECPPIHITARR 435
DB 421 LNMVPTLTFKKN 433

RESULT 11

Q8PF59 PRELIMINARY; PRT; 503 AA.
AC 08PF59;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Probable GTP-binding protein enga.
GN C3033.
OS Escherichia coli O6.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
CX NCBI_TaxID=217992;
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN-06:H1 / CPT073 / ATCC 700928;
RX MEDLINE-22388234; PubMed-12471157;
RA Welch R.A., Burland V., Plunkett G. III, Redford P., Rossch P.,
RA Rasko D., Buckles E.L., Lion S.-R., Boutin A., Hackett J., Stroud D.,
RA Mayhew G.F., Rose D.J., Zhou S., Schwartz D.C., Perna N.T.,
RA Mobley H.L.T., Domeneberg H.S., Blatner F.R.;
RT "Extensive mosaic structure revealed by the complete genome sequence
of uropathogenic Escherichia coli.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:17020-17024 (2002).
DR EMBL; AE016764; AAN81483.1; -.
KW Complete proteome.
SQ SEQUENCE 503 AA; 56573 MW; 7111A9BD763280A CRC64;

Query Match 41.6%; Score 930; DB 16; Length 503;
Best Local Similarity 42.0%; Pred. No. 4.6e-50;

Matches 197; Conservative 91; Mismatches 143; Indels 38; Gaps 4;

QY 4 PIVAVGRPNVSKSTIFNRIYGERVSIYEDTFCVTRDRIRYSSGEMLTGHDENIIDTGCIEI 63
DB 16 PVVALVGRPNVSKSTIFNRLTRTRDALVADFCGLTRDRKYGRAIEGREFICIDGIDG 75
QY 64 GDAPFOTQIRAOAEIAIDEADYIIFMNVNREGLTOSDEMVAQIYKSKRPVLAIVNKNVDN 123
DB 76 TEGCVETRNAQDSLAIEADYVLFMVARAGALPADAIKHLRSKKEPTFLVANKTDG 135
QY 124 MEKRTDYDFYSLGSGEPYPIGSHGGLGLDLDAV----- 160
DB 136 LDQDAVVDYFVLAGELIYPIAASHGRCVLSLLEHVLLPMWEDLAPQEEVDEDAEYWAQF 195
QY 161 ---SHFGSEEDPYEDT--IRLSIIGRPVNGKSSLVNALIGEDRVIVSNVAGTTIRDAID 215
DB 196 EAENECEEEEDDPPOSILPILALVGRPNVSKSTIFNRLIGSEERVVVYDMPTTDSIT 255

QY 216 TETSYDODVYLLDTAGAKKKKGVYESTEKYSVALAKAIERSNVVLVIDAEOGITEOD 275
DB 256 IPMERDREYVLLDTAGVRRRGKITDAVEFVITTLQAIIDANVVLVIDAREGISDOD 315
QY 276 KRAGVAHEOGKAVIYVNMKDPVEKDSKTMKFFEDVREKPEFLDYAOIAFVSAKERTR 335
DB 316 LSLGLFNLNCGRSILYIYVNMKDLDSQVEKQETLDFR--LGTFIDFARVHFISALHGSC 373
QY 336 LRTFPIYNEASENMKKRQVOSTLNEVVDALSMNPTPTDKGRNLNMFYATOVAIEPTF 395
DB 374 VGNFVREZRAIYDSSSTRPGTSMLTRITMAIEDQPPLYNGRKRYKLITAHAGIYNPY 433
QY 396 VFVYNDVLMHFSYKRYLENOIRAFGECPPIHIT-----IARRN 436
DB 434 VTHGNQVKDLPDSYKRYLNMFRKSLDVGSPRIQFREGENPYANRKN 482

RESULT 12

Q8PF02 PRELIMINARY; PRT; 496 AA.
AC 08PF02;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE GTP binding protein.
GN VY10423.
OS Vibrio vulnificus.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
OC Vibrionaceae; Vibrio.
CX NCBI_TaxID=672;
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN-CMCP6;
RA Rhee J.H., Kim S.Y., Chung S.S., Kim J.J., Moon Y.H., Jeong H.,
RA Choy H.E.;
RT "Complete genome sequence of Vibrio vulnificus CMCP6.";
RT Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AE016798; AAO08946.1; -.
KW Complete proteome.
SQ SEQUENCE 496 AA; 55469 MW; 2719F6BA312B7923 CRC64;

Query Match 40.6%; Score 898.5; DB 16; Length 496;
Best Local Similarity 42.4%; Pred. No. 1e-48;

Matches 195; Conservative 84; Mismatches 146; Indels 35; Gaps 4;

QY 4 PIVAVGRPNVSKSTIFNRIYGERVSIYEDTFCVTRDRIRYSSGEMLTGHDENIIDTGCIEI 63
DB 3 PVVALVGRPNVSKSTIFNRLTRSRDALVADFCGLTRDRKYGQAVYGEHDFIVDTGIDG 62
QY 64 GDAPFOTQIRAOAEIAIDEADYIIFMNVNREGLTOSDEMVAQIYKSKRPVLAIVNKNVDN 123
DB 63 SESEVERKMAQSLAIREADYVLFVNDGRALTLTSDAIAHLRIKIEKATMLVANKVDG 122
QY 124 MEMRTDYDFYSLGSGEPYPIGSHGGLGLDLDAVSHF-----GEEBE--DPY 171
DB 123 IDAASADFMQGVDEMVOIAAHGRCGTALIERALDPFDDNLLSANNGSEIETLDMR 182
QY 172 DEPT-----IRLSIIGRPVNGKSSLVNALIGEDRVIVSNVAGTT 210
DB 183 DEDAEQOEYSEEDAEESLRKLDQDPIKILAITIGRRPNVSKSTLTNRLICEERVVYDMPTT 242
QY 211 RDAIDTETSYDODVYLLDTAGAKKKKGVYESTEKYSVALAKAIERSNVVLVIDAREG 270
DB 243 RDSITITMERDREYVLLDTAGVRRRGKITDAVEFVITTLQAIIDANVVLVIDAREG 302
QY 271 IIEODKRVAGAEHOGKAVIYVNMKDPVEKDSKTMKFFEDVREKPEFLDYAOIAFVSA 330
DB 303 ISDODLSLCLFALNAGRSIVLAIVNKNWDCG--DNEVKEVWKELDRRLGPFDFARIHFISA 360
QY 331 KERRLRTLPFYNEASENMKKRQVOSTLNEVVDALSMNPTPTDKGRNLNMFYATOVAIE 390
DB 361 LHGTGVGHLEFESQAEVIRSATIRVGTIVLTRINKMAETEDHPRPVNRGRKRYKLITAHAG 420

```

OY      391 EPPTEVVFVNDVELMHFYSKRYLENDIRAAFGFEGTPIHI 430
        ||| : | | : ||||| | : |||||
Db      421 NPPIVVIHGNOVRELPSYKRYLMYFRKSLDIMGTPIRI 460

```

RESULT 13

ID	09LH58	PRELIMINARY;	PRT;	537	AA.
----	--------	--------------	------	-----	-----

AC Q93AB8;
D7 01-OCT-2000 (TREMBLrel. 15, Created)
D7 01-OCT-2000 (TREMBLrel. 15, Last annotation update)
D7 01-OCT-2002 (TREMBLrel. 22, Last annotation update)
DE GTP-binding protein-like.
OS Arabidopsis thaliana (mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosid II; Brassicales; Brassicaceae; Arabidopsids.
NCBI_TaxID=3702;
[1]
RN SEQUENCE FROM N.A.
RP
RC SPRAIN-Columbe,
RC Kaneko T., Kato T., Sato S., Nakamura Y., Asamizu E., Tabata S.;
RL Submitted (Mar-2000) to the EMBL/GenBank/DBD databases.

Query Match	39.5%	Score 874;	DB 10;	Length 537;
Best Local Similarity	42.3%;	Pred. No. 4e-47;		
Matches 202;	Conservative 81;	Mismatches 143;	Indels 52;	Gaps 10;

[illegible]

Db 391 SLKAMPYISSTALIGHSVDNIIVAAATVOKERSRSTALTNLQVIEENAFSPRTGCG 450

Qy 378 RRLANFATOVAIEPTFVVFVNDVELMHFSYKRYLENOIRAAFCGEPPIHIIARKK 435

451 KRGRIYCTQALRPPTFFVFVNDKLKLSDDYRRIMEKQJLRTAGACGPIPIILMKRSF 508

RESULT 14

ID	09c7c0	PRELIMINARY:	PRT:	659 AA.
AC	09c7c0:			
DT	01-JUN-2001 (TREMBLrel. 17, Created)			
DT	01-JUN-2001 (TREMBLrel. 17, Last sequence update)			
DT	01-OCT-2002 (TREMBLrel. 22, Last annotation update)			
DE	cnpase, putative.			
GN	T21B14.10;			
OS	Arabidopsis thaliana (mouse-ear cress).			
OC	Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;			
OC	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;			
OC	eucosids II; Brassicales; Brassicaceae; Arabidopsi.			
NCBI_Taxid=3702;				

Query Match	. 39.58;	Score 874;	DB 10;	Length 659;
Best Local Similarity	42.38;	Pred. No. 5.4e-47;		
Matches 202;	Conservative 81;	Mismatches 143;	Indels 52;	Gaps 10

```

Oy      6 VAIVGRPNWGKSTIFNRIYIGERVSIVEDTPCVTRDRIRYSGGEMLTHDFNIIDTCGI----- 61
          |||||
          |||||: ||| |||||:| | :|||:
Db      157 VAIVGRPNWGSALFNRLVGNRAIVYDEPCVTRDRRLYKGSYWCQGEFVVVDVTGCVMTYS 216

```



```

Oy 62 -----EIGDAPFOTQIRAOAETADADYIIFVNVYREG 95
Db 217 KSPSCVMEELANSTTIGMCGITPLSREAAIARMSHMKQTAADVSAVIFVYDQAG 276
Oy 96 LTOSDEMAOIL--YKSKRPVLAIVKVDN--MEMRTDVFYSLGFGEPYISGSHGL 152
Db 277 PSQADVEIADWLKRYKSHYIILAVNKCSPKGLMASEFWSGCF--TPIDISLSTGCT 335
Oy 153 GDLDAVVS-----HFGSEEDPYDEDTIRLSIGRPVGSLSVNAIIGEDRVIV 203
Db 336 GELDLVSGGLKLEIMENIEESEEEN--IPAAIIGRPVGSLSVNAIIGEDRVIV 392
Oy 204 SNAAGTTRADITERS--YGCODYVLIOTAGMKKCKYVES--TEKISVLRALKALERSN 259
Db 393 SPVSGTTRADIDELFGPDKGKFRILIDPAGIKKSSVASGSGSTEASVNRARAIERSD 452
Oy 260 VLVVVIDAEGGIIEDDKRVAGYAHGOKRAVYIVVNMKMDV--EKDSTMKKFEDEVRKEFO 318
Db 453 VVALVIEAMACITEODLKIARIEREGKGLVYVNMKMDTIPNKHQETAHAHYEDDVREKLR 512
Oy 319 FLVYQIAYVSAKERTRLRTLPYINLEASNNHKKVOSTLENVYTDIAISM--NPTPTDNG 377
Db 513 SLNMAPIYSTAIGSVNIVYAAATVOKERSRLSTAILNQVIRCAVAFKSPPTRG 572
Oy 378 RLNVFATQVAIEPPTFVYVNDVLAHFSYKRYLENOIRPAFGFPGCTPIHIAKR 435
Db 573 KRCRVYCTOAIIRPTFVFVFNDAKLFSDTYRRYMERQKRTDAGFAGTPIRLMRSR 630

```

RESULT 15

```

O8EC36 PRELIMINARY: PRT: 487 AA.

```

```

AC O8EC36 01-MAR-2003 (Tremblrel. 23, Created)
DT 01-MAR-2003 (Tremblrel. 23, Last sequence update)
DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)
DE GTP-binding protein Enga.
GN ENG4 OR SO3308.
OS Shewanella oneidensis.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Alteromonadales;
OC Alteromonadaceae; Shewanella.
OX NCBI_TaxID=70863;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NR-1;
RX MEDLINE=2297686; Pubmed=12368813;
RA Heidelberg J.F., Paulsen I.T., Nelson K.E., Gaidos E.J., Nelson W.C.,
RA Read T.D., Eisen J.A., Seshadri R., Ward N., Neche B., Clayton R.A.,
RA Meyer T., Tsaplin A., Scott J., Beanan M., Brinkac L., Daugherty S.,
RA DeBoy R.T., Dodson R.J., Durkin A.S., Haft D.H., Kolonay J.F.,
RA Madupu R., Peterson J.D., Umayam L.A., White O., Wolf A.M.,
RA Vamathevan J., Weidman J., Imprial M., Lee K., Berry K., Lee C.,
RA Feldblyum T.V., Smith H.O., Venter J.C., Nelson K.H., Fraser C.M.;
RT Genome sequence of the dissimilatory metal ion-reducing bacterium
RL Nat. Biotechnol. 20:1118-1123(2002).
DR EMBL: AB015769; AAN56306.1;
DR TIGR: SO3308;
KW Complete proteome.
SQ SEQUENCE 487 AA; 54531 MW; B75F8C6CA1BF9C20 CRC64;

```

Query Match 38.8%; Score 859; DB 16; Length 487;

Best Local Similarity 39.98; Pred. No. 3; 1e-46;

Matches 180; Conservative 98; Mismatches 147; Indels 26; Gaps 4;

```

Oy 4 PIVAVIGRPVNGSKSTFNRIIVGERSIVEDTFCVTRDRIRYSSGEMTLHDFNIIIDTGGIEI 63
Db 3 PVALVIGRPVNGSKSTFNRLTRTADALVADFGILTRDRKRCRAFLVSGYEFTIVDTGGIDG 62
Oy 64 GDAPFOTQIRAOAETADADYIIFVNVYREGTQSDMAVNOIYKSKRPVLAIVKVDN 123

```

```

Db 63 TEESIEFKMAEQSLAAIEADAVVLFMTDRAGLTADLSINQHLRSNQKTTFFVANKIDG 122
Oy 124 MEANTDVIDYDPSYSLGFGEPYISGSHGLGDLDAVVSF-----GEEF---EDPY 171
Db 123 IDADSACAEFWSLGLLEVYQMAAAGRGVTNNIEYALPVYEAAGIERGQEEVEDEKRY 182
Oy 172 DEE-----TIRLSIGRPVGSLSVNAIIGEDRVIVSNVAGTTRADITERS 219
Db 183 TEFEAEAEOKRLQDLPIKILAIIGKPNVGSSTLTNNILCEERVVYVDEPGTTSDSIVIME 242
Oy 220 YGODYVLIOTAGMKKCKYVES--TEKISVLRALKALERSN--NPTPTDNG 279
Db 243 RDGREYIYIDTACVRRRSKVHEIYKFSYIKLAVEDAVYVLIIDARCGVAFEDGLGL 302
Oy 280 GYAHGOKRAVYIVVNMKMDTVERKDSKTKMKFEDEVRKEFOPLDYAQIAFVSAKERTRLRTL 339
Db 303 GFALNAGRALYIVVNMKMDIDGIR--DRVSELDORRIGFTDFARHIFISALHGTGVCHL 360
Oy 340 FPYINLEASNNHKKVOSTLENVYTDIAISMNPTPTDNGKRLNVFATQVAIEPPTFVYV 399
Db 361 FESIEEAYDSATRRVSTSLMTRIMQMSQDDHQPLVNGRRAVKLYKAHAGYNPPIYVING 420
Oy 400 NDVLEHFSYKRYLENOIRPAFGFPGCTPIHIAKR 430
Db 421 NOVSKLPDSYKRYMNYFRSLKLVYGTPIQL 451

```

Search completed: September 10, 2003, 00:35:04
Job time : 105 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 12, 2003, 17:16:39 : Search time 129.915 Seconds
(without alignments)
8041.284 Million cell updates/sec

Title: US-09-815-242-1463

Perfect score: 387

Sequence: 1 gatctctctctctctccca.....ttcacctgaagaataaac 387

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2552756 segs, 1349719017 residues

5105512

Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N_Geneseq_19Jun03:*

1: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT.*
2: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT.*
3: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT.*
4: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT.*
5: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT.*
6: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT.*
7: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT.*
8: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT.*
9: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT.*
10: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT.*
11: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT.*
12: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT.*
13: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT.*
14: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT.*
15: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT.*
16: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT.*
17: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT.*
18: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT.*
19: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.*
20: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.*
21: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.*
22: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.*
23: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.*
24: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT.*
25: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	387	100.0	387	23	AA548886
2	387	100.0	1311	22	AA548886
3	387	100.0	1311	23	AA548886
4	387.8	99.2	1305	23	AA548886
5	383.8	99.2	1311	23	AA548886
6	327	84.5	3621	18	AA548886
7	295	76.2	372	23	AA548886
8	263.2	67.8	1332	24	AA548886

9	262.2	67.8	3269	22	AA548886
10	234	60.5	298	23	AA548886
11	234	60.5	298	23	AA548886
12	205.8	53.2	319630	24	AB067194
13	205.8	53.2	2944528	24	AB067194
14	205.8	53.2	3011208	24	AB067194
15	174.4	45.1	1311	24	AB067194
16	162.6	42.0	1308	24	AB067194
17	162.6	42.0	2355589	24	AB067194
18	160.4	41.4	1308	24	AB067194
19	160.4	41.4	2155561	24	AB067194
20	153	39.5	1311	21	AA548886
21	151.4	39.1	1308	25	AB067194
22	151.4	39.1	1311	21	AA548886
23	151.4	39.1	5066	19	AA548886
24	151.4	39.1	2162598	25	AA548886
25	146	37.7	245	23	AA548886
26	101	26.1	960	22	AA548886
27	89.4	22.1	1512	23	AA548886
28	89.4	22.1	1574	22	AA548886
29	86	22.2	157	23	AA548886
30	84.4	21.8	1512	22	AA548886
31	84.4	21.8	1512	22	AA548886
32	84.4	21.8	1515	23	AA548886
33	83.8	21.7	1473	23	AA548886
34	80.6	20.8	640681	24	AA548886
35	80.2	20.7	1830121	17	AA548886
36	73.2	18.9	1370	19	AA548886
37	73	18.9	33140	22	AA548886
38	69.8	18.0	1146	25	AA548886
39	69.8	18.0	1557	22	AA548886
40	69.8	18.0	349980	22	AA548886
41	69.8	18.0	349980	22	AA548886
42	64	16.5	25360	22	AA548886
43	64	16.5	29736	22	AA548886
44	62	16.0	580073	18	AA548886
45	61.4	15.9	78845	21	AA548886

ALIGNMENTS

RESULT 1	AA548886	standard; DNA; 387 BP.
AA548886		
AC	AA548886	
XX		
DT	13-FEB-2002	(first entry)
XX		
DE	Staphylococcus aureus cellular proliferation inhibitory sequence #110.	
XX		
KW	antisense; ss; prokaryotic cellular proliferation;	
KW	antibiotic; antibacterial; drug design.	
XX		
OS	Staphylococcus aureus.	
XX		
PN	WO200170955-A2.	
XX		
PD	27-SEP-2001.	
XX		
PF	21-MAR-2001; 2001WO-US09180.	
XX		
PR	21-MAR-2000; 2000US-191078P.	
PR	23-MAY-2000; 2000US-206848P.	
PR	26-MAY-2000; 2000US-207727P.	
PR	23-OCT-2000; 2000US-242578P.	
PR	27-NOV-2000; 2000US-253625P.	
PR	22-DEC-2000; 2000US-257931P.	
PR	16-FEB-2001; 2001US-269308P.	
XX		
PA	(ELIT-) ELITRA PHARM INC.	
XX		

S. epidermidis gen
Staphylococcus aur
Staphylococcus aur
Listeria innocua c
Listeria monocytog
Listeria innocua D
Bacillus lichenifo
Streptococcus poly
genomic sequence o
Streptococcus poly
Streptococcus poly
ypHC gene of strep
S. pneumoniae type
Streptococcus pneu
Streptococcus pneu
Streptococcus pneu
Staphylococcus aur
S. epidermidis ope
E. coli DNA for ce
DNA encoding novel
Staphylococcus aur
Haemophilus influe
DNA encoding haemo
Haemophilus influe
Salmonella typhi D
Buchnera sp. genom
Haemophilus influe
S. pneumoniae derl
Genomic fragment #
C. glutamicum derl
C. glutamicum codin
C. glutamicum codin
S. spinoza DNA fire
S. spinoza DNA fire
Mycoplasma genital
N. meningitidis pa

OY 301 TTGGATGTCATCACCATTTCATACACCTGTATCATATATTTGAATCATGTG 360
 DB 208 TTGGATGTCATCACCATTTCATACACCTGTATCATATATTTGAATCATGTG 149
 OY 361 TTACCATTCACCTGAGAAATTAATAC 387
 DB 148 TTACCATTCACCTGAGAAATTAATAC 122

RESULT 3

AAS54997/c
 ID AAS54997 standard; DNA: 1311 BP.

AAS54997;
 13-FEB-2002 (first entry)

XX Staphylococcus aureus DNA for cellular proliferation protein #1309.

DE Staphylococcus aureus DNA for cellular proliferation protein #1309.

XX Antisense: ds; prokaryotic cellular proliferation gene;
 KM antibiotic; antibacterial; drug design.

OS Staphylococcus aureus.

XX WO200170955-A2.

PD 27-SEP-2001.

XX 21-MAR-2001; 2001WO-US09180.

PR 21-MAR-2000; 2000US-191078P.

PR 23-MAY-2000; 2000US-206848P.

PR 26-MAY-2000; 2000US-207727P.

PR 23-OCT-2000; 2000US-242578P.

PR 27-NOV-2000; 2000US-253625P.

PR 22-DEC-2000; 2000US-257931P.

PR 16-FEB-2001; 2001US-269308P.

XX (ELIT-) ELITRA PHARM INC.

XX Haselbeck R, Ohlgen KL, Zyckind JW, Wall D, Trawick JD, Carr GJ;
 PI Yamamoto RT, Xu HH;
 DR WPI: 2001-611495/70.
 XX P-PSDB: AAU37138.

XX New polynucleotides for the identification and development of
 PT antibiotics, comprise sequences of antisense nucleic acids -
 PS Claim 27; Seq ID No 8634; 511pp; English.

XX The invention relates to antisense inhibitors of genes essential to
 CC prokaryotic cellular proliferation, their use in identifying the
 CC genes, their use in the discovery of novel antibiotics, the essential
 CC genes themselves and the encoded proteins. The prokaryotes used are
 CC Escherichia coli, Staphylococcus aureus, Salmonella typhi, Klebsiella
 CC pneumoniae, Pseudomonas aeruginosa and Enterococcus faecalis. The
 CC invention is also useful for the identification of potential new targets
 CC for antibiotic development. The antisense nucleic acids can also be used
 CC to identify proteins used in proliferation, to express these proteins,
 CC and to obtain antibodies capable of binding to the expressed proteins.
 CC The proteins can be used to screen compounds in rational drug discovery
 CC programmes. The antisense nucleic acid sequence is also useful to screen
 CC a wide variety of organisms. The present sequence encodes an
 CC essential prokaryotic cellular proliferation protein.
 CC Note: The sequence data for this patent did not form part
 CC of the printed specification, but was obtained in electronic
 CC format directly from WIPO at
 CC ftp://wipo.int/pub/published_pct_sequences.

XX Sequence 1311 BP: 452 A: 184 C: 278 G: 397 T: 0 other:

Query Match 100.0%; Score 387; DB 23; Length 1311;
 Best Local Similarity 100.0%; Pred. No. 1.8e-98;
 Matches 387; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GATCTCTTCCTCTTACCAAAATGAGAAACACTGCATATACCAAGTACCAAGACCTA 60
 DB 508 GATCTCTTCCTCTTACCAAAATGAGAAACACTGCATATACCAAGTACCAAGACCTA 449
 OY 61 AACCATGAGCCCTGATATCGGATACGGTTCCACCAAACTCTAATGATAGAAATCATACA 120
 DB 448 AACCATGAGCCCTGATATCGGATACGGTTCCACCAAACTCTAATGATAGAAATCATACA 389
 OY 121 CGTCTGACGATTTCCATATATATCTACTTTGTAACGGCTAATAGACGGTTTTTNG 180
 DB 388 CGTCTGACGATTTCCATATATATCTACTTTGTAACGGCTAATAGACGGTTTTTNG 329
 OY 181 ATTTGATATAATTTGACGACGACATTTTCATCGCTTTGTGCAATCCTTACGACGTTAA 240
 DB 328 ATTTGATATAATTTGACGACGACATTTTCATCGCTTTGTGCAATCCTTACGACGTTAA 269
 OY 241 CCATATAAATAATATACATCGCTTCATCTATGCGGATTTTCCGCTGCTAATTTGTG 300
 DB 268 CCATATAAATAATATACATCGCTTCATCTATGCGGATTTTCCGCTGCTAATTTGTG 209
 OY 301 TTGGATGTCATCACCATTTCATACACCTGTATCATATATTTGAATCATGTG 360
 DB 208 TTGGATGTCATCACCATTTCATACACCTGTATCATATATTTGAATCATGTG 149
 OY 361 TTACCATTCACCTGAGAAATTAATAC 387
 DB 148 TTACCATTCACCTGAGAAATTAATAC 122

RESULT 4

AAS51646/c
 ID AAS51646 standard; DNA: 1305 BP.

AAS51646;
 13-FEB-2002 (first entry)

XX Staphylococcus aureus DNA for cellular proliferation protein #63.

DE Staphylococcus aureus DNA for cellular proliferation protein #63.

XX Antisense: ds; prokaryotic cellular proliferation gene;
 KM antibiotic; antibacterial; drug design.

OS Staphylococcus aureus.

XX WO200170955-A2.

PD 27-SEP-2001.

XX 21-MAR-2001; 2001WO-US09180.

PR 21-MAR-2000; 2000US-191078P.

PR 23-MAY-2000; 2000US-206848P.

PR 26-MAY-2000; 2000US-207727P.

PR 23-OCT-2000; 2000US-242578P.

PR 27-NOV-2000; 2000US-253625P.

PR 22-DEC-2000; 2000US-257931P.

PR 16-FEB-2001; 2001US-269308P.

XX (ELIT-) ELITRA PHARM INC.

XX Haselbeck R, Ohlgen KL, Zyckind JW, Wall D, Trawick JD, Carr GJ;
 PI Yamamoto RT, Xu HH;
 DR WPI: 2001-611495/70.
 XX P-PSDB: AAU37187.

XX New polynucleotides for the identification and development of
 PT antibiotics, comprise sequences of antisense nucleic acids -
 PS Claim 27; Seq ID No 8634; 511pp; English.

PS Claim 27; Seq ID No 4228; 511pp; English.

CC The invention relates to antisense inhibitors of genes essential to
CC prokaryotic cellular proliferation, their use in identifying the
CC genes, their use in the discovery of novel antibiotics, the essential
CC genes themselves and the encoded proteins. The prokaryotes used are
CC Escherichia coli, Staphylococcus aureus, Salmonella typhi, Klebsiella
CC pneumoniae, Pseudomonas aeruginosa and Enterococcus faecalis. The
CC invention is also useful for the identification of potential new targets
CC for antibiotic development. The antisense nucleic acids can also be used
CC to identify proteins used in proliferation, to express these proteins,
CC and to obtain antibodies capable of binding to the expressed proteins.
CC The proteins can be used to screen compounds in rational drug discovery
CC programmes. The antisense nucleic acid sequence is also useful to screen
CC for homologous nucleic acids which are required for cell proliferation in
CC a wide variety of organisms. The present sequence encodes an
CC essential prokaryotic cellular proliferation protein.
CC Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic
CC format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.

CC Sequence 1305 BP; 447 A; 181 C; 277 G; 400 T; 0 other;

Query Match 99.2%; Score 383.8; DB 23; Length 1305;
Best Local Similarity 99.5%; Pred. No. 1.4e-97;
Matches 385; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GATCTTCTCTCTCTCACCACCAATGAGAAACAGTCGATCTAACAAGTACACAGACCTA 60
DB 508 GATCTTCTCTCTCTCACCACCAATGAGAAACAGTCGATCTAACAAGTACACAGACCTA 449
QY 61 AACCATGAGCCGATGATGCGATGCGATGCGATGCGATGCGATGCGATGCGATGCGAT 120
DB 448 AACCATGAGCCGATGATGCGATGCGATGCGATGCGATGCGATGCGATGCGATGCGAT 389
QY 121 CGTCTGATGCGATGATGCGATGCGATGCGATGCGATGCGATGCGATGCGATGCGAT 180
DB 388 CGTCTGATGCGATGATGCGATGCGATGCGATGCGATGCGATGCGATGCGATGCGAT 329
QY 181 ATTTGATTAATAATTTGAGCGACGATTTGATGCGATTTGATGCGATTTGATGCGATTTA 240
DB 328 ATTTGATTAATAATTTGAGCGACGATTTGATGCGATTTGATGCGATTTGATGCGATTTA 269
QY 241 CCATATAAATAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 300
DB 268 CCATATAAATAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 209
QY 301 TTTGGAATGTCATCACCACCAATTTCAATACCACTGATCAATTAATTAATTAATTAAT 360
DB 208 TTTGGAATGTCATCACCACCAATTTCAATACCACTGATCAATTAATTAATTAATTAAT 149
QY 361 TTAACCATTCACCTGAGAAATTAATAC 387
DB 148 TTAACCATTCACCTGAGAAATTAATAC 122

RESULT 5
AAS54865/c
ID AAS54865 standard; DNA: 1311 BP.

XX AAS54865;
XX
XX 13-FEB-2002 (first entry)
DE Staphylococcus aureus DNA for cellular proliferation protein #1177.
KW Antisense; ds: prokaryotic cellular proliferation gene;
KW antibiotic; antibacterial; drug design.
XX Staphylococcus aureus.
XX
PN W0200170955-A2.

XX 27-SEP-2001.

PD 21-MAR-2001; 2001MO-US09180.

XX 21-MAR-2001; 2000US-191078P.

PR 23-MAY-2000; 2000US-206848P.

PR 26-MAY-2000; 2000US-207727P.

PR 23-OCT-2000; 2000US-242578P.

PR 27-NOV-2000; 2000US-253625P.

PR 22-DEC-2000; 2000US-257931P.

PR 16-FEB-2001; 2001US-269308P.

PA (ELIT-) ELITRA PHARM INC.

XX Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Travack JD, Carr GU;

PI Yamamoto RT, Xu HH;

DR WPI: 2001-611495/70.

XX P-PSDB; AAU37006.

XX New polynucleotides for the identification and development of

PT antibiotics, comprise sequences of antisense nucleic acids -

XX Claim 27; Seq ID No 8502; 511pp; English.

CC The invention relates to antisense inhibitors of genes essential to
CC prokaryotic cellular proliferation, their use in identifying the
CC genes, their use in the discovery of novel antibiotics, the essential
CC genes themselves and the encoded proteins. The prokaryotes used are
CC Escherichia coli, Staphylococcus aureus, Salmonella typhi, Klebsiella
CC pneumoniae, Pseudomonas aeruginosa and Enterococcus faecalis. The
CC invention is also useful for the identification of potential new targets
CC for antibiotic development. The antisense nucleic acids can also be used
CC to identify proteins used in proliferation, to express these proteins,
CC and to obtain antibodies capable of binding to the expressed proteins.
CC The proteins can be used to screen compounds in rational drug discovery
CC programmes. The antisense nucleic acid sequence is also useful to screen
CC for homologous nucleic acids which are required for cell proliferation in
CC a wide variety of organisms. The present sequence encodes an
CC essential prokaryotic cellular proliferation protein.
CC Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic
CC format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.

CC Sequence 1311 BP; 451 A; 181 C; 277 G; 402 T; 0 other;

Query Match 99.2%; Score 383.8; DB 23; Length 1311;
Best Local Similarity 99.5%; Pred. No. 1.4e-97;
Matches 385; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GATCTTCTCTCTCTCACCACCAATGAGAAACAGTCGATCTAACAAGTACACAGACCTA 60
DB 508 GATCTTCTCTCTCTCACCACCAATGAGAAACAGTCGATCTAACAAGTACACAGACCTA 449
QY 61 AACCATGAGCCGATGATGCGATGCGATGCGATGCGATGCGATGCGATGCGATGCGAT 120
DB 448 AACCATGAGCCGATGATGCGATGCGATGCGATGCGATGCGATGCGATGCGATGCGAT 389
QY 121 CGTCTGATGCGATGATGCGATGCGATGCGATGCGATGCGATGCGATGCGATGCGAT 180
DB 388 CGTCTGATGCGATGATGCGATGCGATGCGATGCGATGCGATGCGATGCGATGCGAT 329
QY 181 ATTTGATTAATAATTTGAGCGACGATTTGATGCGATTTGATGCGATTTGATGCGATTTA 240
DB 328 ATTTGATTAATAATTTGAGCGACGATTTGATGCGATTTGATGCGATTTGATGCGATTTA 269
QY 241 CCATATAAATAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 300
DB 268 CCATATAAATAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 209
QY 301 TTTGGAATGTCATCACCACCAATTTCAATACCACTGATCAATTAATTAATTAATTAAT 360

```

Db      208 TTTGGAAATGGTGCATCCCAATTTCATACCACTGTATCAATTAATTTGAAATCATGTG 149
Qy      361 TTAACCATTCACCGTAGAATATAATAC 387
Db      148 TTACCACTCACCGTAGAATATAATAC 122

RESULT 6
AAV74669/C
ID      AAV74669 standard; DNA; 3621 BP.
XX
XX      AAV74669;
XX
XX      16-MAR-1999 (first entry)
XX
XX      Staphylococcus aureus contig SEQ ID #358.
XX
XX      Computer readable medium; vaccine; S.aureus infection; immunodetection;
XX      cellulitis; eyelid infection; food poisoning; osteomyelitis; therapy;
XX      skin infection; surgical wound infection; scalded skin syndrome;
XX      toxic shock syndrome; ds.
XX
XX      Staphylococcus aureus.
XX
XX      Key      Location/Qualifiers
XX      misc_feature 481..540
XX      /tag=8
XX      /note="these bases represent a line of missing text in
XX      the sequence listing in the specification. They
XX      are included to maintain the nucleotide numbering
XX      given in the specification for this DNA sequence"
XX
XX      misc_feature 2281..2340
XX      /tag=b
XX      /note="these bases represent a line of missing text in
XX      the sequence listing in the specification. They
XX      are included to maintain the nucleotide numbering
XX      given in the specification for this DNA sequence"
XX
XX      EP786519-A2.
XX
XX      30-JUL-1997.
XX
XX      07-JAN-1997; 97EP-0100117.
XX
XX      05-JAN-1996; 96US-0009861.
XX
XX      (HUMA-) HUMAN GENOME SCI INC.
XX
XX      Barash SC, Choi GH, Dillon PJ, Fannon MR, Kunsch CA;
XX      Rosen CA;
XX
XX      WPI; 1997-374922/35.
XX
XX      Polynucleotide(s) and proteins derived from Staphylococcus aureus
XX      stored on computer readable medium and used in the production of
XX      anti-S.aureus vaccines
XX
XX      Claim 1: Page 1241-1243; 3271pp; English.
XX
XX      This sequence represents one of 5191 Staphylococcus aureus DNA sequences
XX      of the invention. The DNA sequences are recorded on a computer readable
XX      medium, preferably selected from a floppy or hard disk, random access
XX      memory (RAM), read-only memory (ROM) or CD-ROM. Homology searches using
XX      the S.aureus DNA sequences allows putative functions to be assigned so
XX      that protein-encoding or regulatory regions of commercial, therapeutic or
XX      industrial importance can be obtained. Specifically, sequences which are
XX      likely to encode antigens have been identified and these polypeptides can
XX      be used in a vaccine composition against S.aureus infection. The
XX      S.aureus in a sample. S.aureus is implicated in numerous human diseases,
XX      including cellulitis, eyelid infections, food poisoning, osteomyelitis,
XX      skin and surgical wound infections, scalded skin syndrome, toxic shock

```

```

CC      syndrome, etc. Organisms transformed with the DNA sequences can be used
CC      for recombinant production of the polypeptides. The new DNA sequences
CC      (and their fragments) are useful as primers or probes for isolating
CC      homologues of any of the S.aureus DNA sequences contained on the
CC      computer readable medium.
XX
XX      SQ      Sequence 3621 BP; 1279 A; 442 C; 692 G; 1085 T; 123 other:
XX
XX      Query Match      84.5%; Score 327; DB 18; Length 3621;
XX      Best Local Similarity 84.5%; Pred. No. 1,6e-81;
XX      Matches 327; Conservative 0; Mismatches 60; Indels 0; Gaps 0;
XX
Qy      1 GATCTTCCTCCCTTCACCAAAATGAGAAACAACATGCATCTAACAAGTCCACCAAGCTTA 60
Db      2637 GATCTTCCTCCCTTCACCAAAATGAGAAACAACATGCATCTAACAAGTCCACCAAGCTTA 2578
Qy      61 AACCATGTGACCCGATATCGGATACGCTTCACCAAAATGCTAATGATGAATACATACA 120
Db      2577 AACCATGTGACCCGATATCGGATACGCTTCACCAAAATGCTAATGATGAATACATACA 2518
Qy      121 CGTCGTACGCGATTTCATATATCTACTTGTAAACCGTAATAGACCGCTTTTATAG 180
Db      2517 CGTCGTACGCGATTTCATATATCTACTTGTAAACCGTAATAGACCGCTTTTATAG 2458
Qy      181 ATTTGTAATAAATTTGACGACGACATTCATCGCTTGTGTCAATCCTTCAGCAGCTTAA 240
Db      2457 ATTTGTAATAAATTTGACGACGACATTCATCGCTTGTGTCAATCCTTCAGCAGCTTAA 2398
Qy      241 CCATATAAATAATACATCCGCTTCATCTATGCGCATTTCTGCTGCGCTTAATTTGTG 300
Db      2397 CCATATAAATAATACATCCGCTTCATCTATGCGCATTTCTGCTGCGCTTAATTTNN 2338
Qy      301 TTTGGAATGGTGATCATCAATTCATATACACCTGTATCAATATTAATGAAATCATGTG 360
Db      2337 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNCTG 2278
Qy      361 TTAACCATTCACCGTAGAATATAATAC 387
Db      2277 TTAACCATTCACCGTAGAATATAATAC 2251

RESULT 7
AAS50706
ID      AAS50706 standard; DNA; 372 BP.
XX
XX      AAS50706;
XX
XX      13-FEB-2002 (first entry)
XX
XX      Staphylococcus aureus cellular proliferation inhibitory sequence #1930.
XX
XX      Antisense; ss; prokaryotic cellular proliferation;
XX      antibiotic; antibacterial; drug design.
XX
XX      Staphylococcus aureus.
XX
XX      WO200170955-A2.
XX
XX      27-SEP-2001.
XX
XX      21-MAR-2001; 2001WO-US09180.
XX
XX      21-MAR-2000; 2000US-191078P.
XX      23-MAY-2000; 2000US-206848P.
XX      26-MAY-2000; 2000US-207272P.
XX      23-OCT-2000; 2000US-242578P.
XX      27-NOV-2000; 2000US-253625P.
XX      22-DEC-2000; 2000US-259331P.
XX      16-FEB-2001; 2001US-269308P.
XX
XX      (ELITR-) ELITRA PHARM INC.
XX
XX      Haselbeck R, Ohlsen KU, Zyskind JW, Wall D, Travlich JD, Carr GJ;

```

PI Yamamoto RT, Xu HH;

DR WPI: 2001-611495/70.

PT New polynucleotides for the identification and development of
antibiotics, comprise sequences of antisense nucleic acids -

PS Claim 1: Seq ID NO 3283; 511pp; English.

xx The invention relates to antisense inhibitors of genes essential to
CC prokaryotic cellular proliferation, their use in identifying the
CC genes, their use in the discovery of novel antibiotics, the essential
CC themselves and the encoded proteins. The prokaryotes used are
CC *Escherichia coli*, *Staphylococcus aureus*, *Salmonella typhi*, *Klebsiella*
CC *pneumoniae*, *Pseudomonas aeruginosa* and *Enterococcus faecalis*. The
CC invention is also useful for the identification of potential new targets
CC for antibiotic development. The antisense nucleic acids can also be used
CC to identify proteins used in proliferation, to express these proteins,
CC and to obtain antibodies capable of binding to the expressed proteins.
CC The proteins can be used to screen compounds in rational drug discovery
CC programmes. The antisense nucleic acid sequence is also useful to screen
CC for homologous nucleic acids which are required for cell proliferation in
CC a wide variety of organisms. The present sequence is an antisense
CC oligonucleotide of the invention.
CC Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic
CC format directly from WIPO at
CC ftp.wipo.int/pub/published_pcl_sequences.

xx Sequence 372 BP; 113 A; 83 C; 51 G; 125 T; 0 other;

SO Query Match 76.2%; Score 295; DB 23; Length 372;

Best Local Similarity 100.0%; Pred. No. 7e-73;

Matches 295; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 93 CCAATCTTAATGATAGAAATCATACAGCTCTGTACGATTCATATATCTACTTGG 152
DB 1 CCAATCTTAATGATAGAAATCATACAGCTCTGTACGATTCATATATCTACTTGG 60
OY 153 TTAACCGGTAATAGACGGGTTTTAGATTGTATTAATTTGAGCGACCATTCATCG 212
DB 61 TTAACCGGTAATAGACGGGTTTTAGATTGTATTAATTTGAGCGACCATTCATCG 120
OY 213 CTTTGTGCAATCTTACGACGCTTAACCAATAAATAAATTAACATCCGCTCATCTATG 272
DB 121 CTTTGTGCAATCTTACGACGCTTAACCAATAAATAAATTAACATCCGCTCATCTATG 180
OY 273 GCGATTTCGCGCTGCTTAATTTGTTGGAAATGTCATCACCATTCAATACCA 332
DB 181 GCGATTTCGCGCTGCTTAATTTGTTGGAAATGTCATCACCATTCAATACCA 240
OY 333 CCGTATCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 387
DB 241 CCGTATCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 295

RESULT 8

ABN90883/c
ID ABN90883 standard; DNA: 1332 BP.

xx ABN90883;

xx 24-JUL-2002 (first entry)

xx *Staphylococcus epidermidis* ORF nucleic acid sequence SEQ ID NO:346.

xx *Staphylococcus epidermidis*; open reading frame; ORF; bacterial infection;

xx antibacterial; gene therapy; gene; ds.

xx *Staphylococcus epidermidis*.

xx US6380370-B1.

PD 30-APR-2002.

xx 13-AUG-1998; 98US-0134001.

xx 14-AUG-1997; 97US-055779P.

xx 08-NOV-1997; 97US-064964P.

xx (GENO-) GENOME THERAPEUTICS CORP.

xx Doucette-Stamm LA, Bush D;

xx WPI: 2002-381255/41.

xx P-PSDB: ABP38338.

xx Novel isolated nucleic acid encoding a *Staphylococcus epidermidis*
PT polypeptide, useful for diagnosing and treating bacterial infections -

PS Disclosure: SEQ ID 346; 267pp; English.

CC ABN90538 to ABN93374 represent *Staphylococcus epidermidis* open reading
CC frame (ORF) nucleic acid sequences which encode the amino acid sequences
CC given in ABP35124 to ABP37960. The *S. epidermidis* sequences have
CC antibacterial activity and can be used in gene therapy. The sequences
CC can also be used in the diagnosis and treatment of bacterial infections,
CC particularly *S. epidermidis* infections. The sequences can be used to
CC screen for compounds able to interfere with the *S. epidermidis* life
CC cycle or inhibit *S. epidermidis* infection.
CC N.B. The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from the
CC USPTO web site.

xx Sequence 1332 BP; 465 A; 190 C; 267 G; 410 T; 0 other;

SO Query Match 67.8%; Score 262.2; DB 24; Length 1332;

Best Local Similarity 79.8%; Pred. No. 1.6e-63;

Matches 309; Conservative 0; Mismatches 78; Indels 0; Gaps 0;

OY 1 GATCTTCCTCCCTTCACCAAAATGAGAAACATGCTATCAACATGCACCAAGACCTA 60
DB 529 GATCTTCCTCCCTTCATTTATTAAGTTTCAACAACTGCATCTAGCAATCTCCAGTCTTA 470
OY 61 AACCATGTGACCTGATATCGATAGAGGTTACCAAAATCTTAATAGATAAATCATACA 120
DB 469 ATCCATGTGACCAAGAAATAGATATGATGTCCAAAGCCCAATGAATAGAAATCATACA 410
OY 121 CGTCTGATGCAATTTTCATATATCTACTTGTATTAACCGTAATAGACCGGTTTGTAG 180
DB 409 TATCATTTAGCAATTTCAAGATATCACTTATTACAGAGTAATAGAACAGGTTTCTTAG 350
OY 181 ATTGTATTAATTTGAGAGGACATTTATGCTTTGTGCAATCTTACAGCACTTA 240
DB 349 ATTGTATTAATTTGAGAGGACATTTATGCTTTGTGCAATCTTACAGCACTTA 290
OY 241 CCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 300
DB 289 CCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 230
OY 301 TTTGGAATGTCATACCAATTTCAATACACCTGATCAATTAATTAATTAATTAAT 360
DB 229 TTTGGAATGTCATACCAATTTCAATACACCTGATCAATTAATTAATTAATTAAT 170
OY 361 TTAACCATTCACCTGAGAAATTAATAC 387
DB 169 TTAACCATTCACCTGAGAAATTAATAC 143

RESULT 9

AAH54708
ID AAH54708 standard; DNA: 3269 BP.

xx AAH54708;

xx 03-SEP-2001 (first entry)

Oy	1	GACCTCTCCCTCTCACCACAAATGAGAAACACATGCATCTACAGTCACCAACAGCTA	60
Db	2297	GACCTCTCATCTTTATTAAAGTTTGCAACACTGCATCTGACCAATCTCCAACTCCTA	2356
Oy	61	AACCATGTACCCCTATATGCGATACGGTTCACCAATCTCTATGAAATGAAATATATCA	120
Db	2357	ATGCATCTGAAACCAAAATAGGAGATTCGATCTCCAAAGCCTAAAGATGAAATATATGAA	2416
Oy	121	CGCTGTACGCATTTCCATTAATCACTTTGTTAAACCGTAATACGACCGGTTTTTAG	180
Db	2417	TATCATATTACGCAATTTCAAGATTAATCAACTTTATTCACAGCTAATATACAGGTTCTTTC	2476
Oy	181	ATTGTATTAATTTGACGAGCAACATTCATCGCTTGTGTCAATCTTCAGCACCTTAA	240
Db	2477	ATTATTAAGGCAATTTGTGCGACCAATTCGTCACATTTGTGTAACTCTCTTCTCAAAATTTGA	2536
Oy	241	CCATATAAATAATACATCGCCTTCATCTATGGGAGTTTTCGCCGCGCTCTAATTTGTG	300
Db	2537	CCATATAAATGATGACATCTCTCTTCAATATGCTATATTCGGCTGTGCACGAGATTTAG	2596

SQ Sequence 298 BP; 91 A; 69 C; 43 G; 95 T; 0 other;

Query Match 60.5%; Score 234; DB 23; Length 298;
 Best Local Similarity 100.0%; Pred. No. 7.9e-56;
 Matches 234; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GATCTTCCTTCCTTCACCAAAATGAGAAACATGTCATTAACAAGTCACCAAGACCTA 60
 DB 65 GATCTTCCTTCCTTCACCAAAATGAGAAACATGTCATTAACAAGTCACCAAGACCTA 124
 OY 61 AACCATGTGACCGCTGATTCGATTCGATTCACCAAAATGAGAAACATGTCATTAACAAGTCACCAAGACCTA 120
 DB 125 AACCATGTGACCGCTGATTCGATTCGATTCACCAAAATGAGAAACATGTCATTAACAAGTCACCAAGACCTA 184
 OY 121 CGTCTGTACGCAATTCATTAATTAATCTACTTTGTTAACCGGTAATACGACCGGTTTTAG 180
 DB 185 CGTCTGTACGCAATTCATTAATTAATCTACTTTGTTAACCGGTAATACGACCGGTTTTAG 244
 OY 181 ATTGTATATAAATTGAGCGACCATTCATCGCTTTGTGTCATTCCTTCACGCA 234
 DB 245 ATTGTATATAAATTGAGCGACCATTCATCGCTTTGTGTCATTCCTTCACGCA 298

RESULT 11

AAS50723 ID AAS50723 standard: DNA: 298 BP.

AC AAS50723:

DT 13-FEB-2002 (first entry)

DE Staphylococcus aureus cellular proliferation inhibitory sequence #1947.

KM Antisense; ss: prokaryotic cellular proliferation;
 KX antiobiotic; antibacterial; drug design.

OS Staphylococcus aureus.

PN W0200170935-A2.

PD 27-SEP-2001.

PR 21-MAR-2001; 2001MO-US09180.

PR 21-MAR-2000; 2000US-191078P.

PR 23-MAY-2000; 2000US-206848P.

PR 26-MAY-2000; 2000US-207127P.

PR 23-OCT-2000; 2000US-243578P.

PR 27-NOV-2000; 2000US-253625P.

PR 22-DEC-2000; 2000US-257931P.

PR 16-FEB-2001; 2001US-269308P.

PA (ELIT-) ELITRA PHARM INC.

PI Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr CJ;
 PI Yamamoto RT, Xu HH;

DR WPI: 2001-611495/70.

XX New polynucleotides for the identification and development of
 PT antibiotics, comprise sequences of antisense nucleic acids -

PS Claim 1: Seq ID No 3300; 511pp; English.

CC The invention relates to antisense inhibitors of genes essential to
 CC prokaryotic cellular proliferation, their use in identifying the
 CC genes, their use in the discovery of novel antibiotics, the essential
 CC genes themselves and the encoded proteins. The prokaryotes used are
 CC *Escherichia coli*, *Staphylococcus aureus*, *Salmonella typhi*, *Klebsiella*
 CC *pneumoniae*, *Pseudomonas aeruginosa* and *Enterococcus faecalis*. The
 CC invention is also useful for the identification of potential new targets
 CC for antibiotic development. The antisense nucleic acids can also be used
 CC to identify proteins used in proliferation, to express these proteins,
 CC and to obtain antibodies capable of binding to the expressed proteins.
 CC The proteins can be used to screen compounds in rational drug discovery

CC programmes. The antisense nucleic acid sequence is also useful to screen
 CC for homologous nucleic acids which are required for cell proliferation in
 CC a wide variety of organisms. The present sequence is an antisense
 CC oligonucleotide of the invention.
 CC Note: The sequence data for this patent did not form part
 CC of the printed specification, but was obtained in electronic
 CC format directly from Mipo at
 CC ftp.vipo.int/pub/published_Pct_sequences.

SO Sequence 298 BP; 91 A; 69 C; 43 G; 95 T; 0 other;

Query Match 60.5%; Score 234; DB 23; Length 298;
 Best Local Similarity 100.0%; Pred. No. 7.9e-56;
 Matches 234; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GATCTTCCTTCCTTCACCAAAATGAGAAACATGTCATTAACAAGTCACCAAGACCTA 60
 DB 65 GATCTTCCTTCCTTCACCAAAATGAGAAACATGTCATTAACAAGTCACCAAGACCTA 124
 OY 61 AACCATGTGACCGCTGATTCGATTCGATTCACCAAAATGAGAAACATGTCATTAACAAGTCACCAAGACCTA 120
 DB 125 AACCATGTGACCGCTGATTCGATTCGATTCACCAAAATGAGAAACATGTCATTAACAAGTCACCAAGACCTA 184
 OY 121 CGTCTGTACGCAATTCATTAATTAATCTACTTTGTTAACCGGTAATACGACCGGTTTTAG 180
 DB 185 CGTCTGTACGCAATTCATTAATTAATCTACTTTGTTAACCGGTAATACGACCGGTTTTAG 244
 OY 181 ATTGTATATAAATTGAGCGACCATTCATCGCTTTGTGTCATTCCTTCACGCA 234
 DB 245 ATTGTATATAAATTGAGCGACCATTCATCGCTTTGTGTCATTCCTTCACGCA 298

RESULT 12

AB067194/c ID AB067194 standard: DNA: 319630 BP.

AC AB067194:

DT 29-AUG-2002 (first entry)

DE *Listeria innocua* contig DNA sequence #7.

KM Antibacterial; *Listeria*; food contamination; mutational analysis;
 KX infection; ds.

OS *Listeria innocua*.

PN W0200228891-A2.

PD 11-APR-2002.

PR 04-OCT-2001; 2001MO-FR03061.

PR 04-OCT-2000; 2000FR-0012697.

PA (INSP) INST PASTEUR.

PA (CNRS) CNRS CENT NAT RECH SCI.

PI Kunst F, Glaser P;

DR WPI: 2002-332479/37.

XX New genomic sequences from *Listeria* species, useful for detection,
 PT treatment and prevention of infection, also related polypeptides,
 PT antibodies and modulators -

PS Claim 5: SEQ ID 7; 180pp; French.

CC The present invention relates to nucleic acid sequences
 CC (AB067198-AB071213) from *Listeria* sp. The sequences are useful as probes
 CC and primers for identification and/or detection of *Listeria* (e.g. as
 CC contaminants in foods, or mutational analysis) and for analysis of
 CC gene expression. Proteins encoded by the nucleic acid sequences can be

Sat, Sep 13 12:47:59 2003

us-09-815-242-1463.rng

Page 11

[illegible]

Search completed: September 12, 2003, 17:32:09
Job time : 138.915 secs

Sat Sep 13 12:47:59 2003

us-09-815-242-1463.rge

Page 1

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 12, 2003, 17:17:09 : Search time 1580.02 Seconds

(without alignments)
10020.134 Million cell updates/sec

Title: US-09-815-242-1463

Perfect score: 387

Sequence: 1 gatctcttccttccttcacca.....ttcacctgaagaataaac 387

IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2888711 seqs, 2045481386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database :

GenBank:*

1: gb_ba:*

2: gb_hlg:*

3: gb_in:*

4: gb_om:*

5: gb_ov:*

6: gb_pat:*

7: gb_ph:*

8: gb_pl:*

9: gb_pr:*

10: gb_ro:*

11: gb_sts:*

12: gb_sy:*

13: gb_un:*

14: gb_vl:*

15: em_ba:*

16: em_fun:*

17: em_hum:*

18: em_in:*

19: em_mu:*

20: em_om:*

21: em_or:*

22: em_ov:*

23: em_pat:*

24: em_ph:*

25: em_pl:*

26: em_ro:*

27: em_sts:*

28: em_un:*

29: em_vl:*

30: em_hlg_hum:*

31: em_hlg_inv:*

32: em_hlg_other:*

33: em_hlg_mus:*

34: em_hlg_pln:*

35: em_hlg_rod:*

36: em_hlg_mam:*

37: em_hlg_vrt:*

38: em_sy:*

39: em_higo_hum:*

40: em_higo_mus:*

41: em_higo_other:*

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	387	100.0	1308	6	AX622668
C 2	387	100.0	301550	1	AP003134
C 3	387	100.0	333750	1	AP004827
C 4	387	100.0	346900	1	AP003362
C 5	262.2	67.8	3269	1	AF270032
C 6	262.2	67.8	3269	6	AX145350
C 7	262.2	67.8	300892	1	AE016747
C 8	210.6	54.4	290117	1	AE017028
C 9	209	54.0	304680	1	AE017002
C 10	205.8	53.2	313450	1	AL596170
C 11	205.8	53.2	319630	6	AX413016
C 12	205.8	53.2	347050	1	AL591981
C 13	205.8	53.2	349980	6	AX417046
C 14	205.8	53.2	349980	6	AX641670
C 15	198.4	51.3	24887	1	BACSERA
C 16	198.4	51.3	213680	1	BSUB0012
C 17	191.4	49.5	302173	1	AE016951
C 18	188.2	48.6	300550	1	AP004599
C 19	185	47.8	300550	1	AP001512
C 20	182.4	47.1	302050	1	AL935257
C 21	174.4	45.1	1311	6	AX433884
C 22	162.6	42.0	11071	1	AE006309
C 23	162.6	42.0	12434	1	AE006498
C 24	162	41.9	1311	6	AX607165
C 25	162	41.9	44145	6	AX602195
C 26	162	41.9	174050	1	SAG766852
C 27	160.4	41.4	20601	1	AE014265
C 28	159.4	41.2	52276	1	AP005146
C 29	157.8	40.8	12370	1	AE009978
C 30	157.8	40.8	12370	1	AE009978
C 31	153	39.5	12340	1	AE008523
C 32	151.4	39.1	1308	6	AX570314
C 33	151.4	39.1	5066	6	AR218847
C 34	151.4	39.1	10310	1	BD003759
C 35	151.4	39.1	10310	2	SPNEU1902
C 36	151.4	39.1	151947	1	AE007464
C 37	151.4	39.1	349980	6	AX571764
C 38	151.4	39.1	349980	6	AX571765
C 39	149.8	38.7	13860	1	AE015016
C 40	148.2	38.3	3737	1	AB016077
C 41	139.4	36.0	301278	1	AE015939
C 42	125.2	32.4	296750	1	AP003191
C 43	117	30.2	10861	1	AE007680
C 44	109.2	28.2	3557	1	AY094626
C 45	101	26.1	960	6	AX144037

ALIGNMENTS

RESULT 1
AX622668/c AX622668 1308 bp DNA linear PAT 20-FEB-2003
LOCUS Sequence 5631 from Patent WO02094868.
DEFINITION AX622668
ACCESSION AX622668
VERSION AX622668.1 GI:28450653
KEYWORDS
SOURCE
ORGANISM
Staphylococcus aureus
Staphylococcus aureus
Bacteria; Firmicutes; Bacillales; Staphylococcus.
REFERENCE
1
Maignanl,V.C., Mora,M.C. and Scarselli,M.C.
TITLE
Staphylococcus aureus proteins and nucleic acids
Patent: WO 02094868-A 5631 28-NOV-2002;
JOURNAL
Chilton Sps (IT)

Prod. No. is the number of results predicted by chance to have a

```

FEATURES
    source
        Location/Qualifiers
            1..1308
                /organism="Staphylococcus aureus"
                /mol_type="genomic DNA"
                /db_xref="taxon:1280"
BASE COUNT      450 a      184 c      278 g      396 t
ORIGIN
Query Match      100.0% Score 387; DB 6; Length 1308:
Best Local Similarity 100.0%: Pct. NO. 2e-77; 0; Indels 0; Gaps 0;
Matches 387; Conservative 0; Mismatches 0;

QY 1 GATCTCTCTCTCTCTACCAAAATGAGAACAACATGTCATTAACAAGTCACCAAGCCTA 60
DB 508 GATCTCTCTCTCTCTACCAAAATGAGAACAACATGTCATTAACAAGTCACCAAGCCTA 449
QY 61 AACCATGTGACCTGATATCGGATAGCGTTCCACCAATCTTAATGATGAAATCATACA 120
DB 448 AACCATGTGACCTGATATCGGATAGCGTTCCACCAATCTTAATGATGAAATCATACA 389
QY 121 CGTCTGTACGCAATTTCCATATTACTTTGTTAAACCGCTAATACGACCGGTTTTTAG 180
DB 388 CGTCTGTACGCAATTTCCATATTACTTTGTTAAACCGCTAATACGACCGGTTTTTAG 329
QY 181 ATTTGTATAAATTTTGACGACGACATTTTCATCGCTTTGTGATCCTGACGACGCTTAA 240
DB 328 ATTTGTATAAATTTTGACGACGACATTTTCATCGCTTTGTGATCCTGACGACGCTTAA 269
QY 241 CCATATAAATAATTAACATCGCTTCATATGCGGATTTGCGCTGCGCTTAATTTGTG 300
DB 266 CCATATAAATAATTAACATCGCTTCATATGCGGATTTGCGCTGCGCTTAATTTGTG 209
QY 301 TTGGATGTGTCATCACCATTTCATTAACACCGCTATCAATATTAATGAAATCATGTG 360
DB 208 TTGGATGTGTCATCACCATTTCATTAACACCGCTATCAATATTAATGAAATCATGTG 149
QY 361 TTAACCATTCACCTGAGGAATAATATAC 387
DB 148 TTAACCATTCACCTGAGGAATAATATAC 122

RESULT 2
LOCUS      AP003134      301550 bp      DNA      linear      BCT 24-APR-2003
DEFINITION Staphylococcus aureus subsp. aureus N315 genomic DNA, complete
ACCESSION  AP003134 BA000018
VERSION     AP003134.2 GI:14349226
KEYWORDS
SOURCE      Staphylococcus aureus subsp. aureus N315
ORGANISM    Staphylococcus aureus subsp. aureus N315
REFERENCE   1
AUTHORS     Kuroda,M., Ohta,T., Uchiyama,I., Baba,T., Yuzawa,H., Kobayashi,I.,
            Cui,L., Oguchi,A., Aoki,K., Nagai,Y., Imai,J., Ito,T., Kanamori,M.,
            Matsunaru,H., Maruyama,A., Murakami,H., Hosoyama,A.,
            Mizutani,O.I.,Y., Takahashi,N.K., Sawano,T., Inoue,R., Kaito,C.,
            Sekimizu,K., Hiraoka,H., Kuwano,S., Goto,S., Yabuzaki,J.,
            Kanehisa,M., Yamashita,A., Oshima,K., Furuya,K., Yoshino,C.,
            Shiba,T., Hattori,M., Ogasawara,N., Hayashi,H. and Hiramatsu,K.
            Whole genome sequencing of methicillin-resistant Staphylococcus
            aureus
JOURNAL     Lancet 357 (9264), 1225-1240 (2001)
MEDLINE     21311952
PUBMED      11418166
REFERENCE   2 (bases 1 to 301550)
AUTHORS     Director-General, Biotechnology Center, Aoki,K., Oguchi,A.,
            Hosoyama,A., Nagai,Y., Kuroda,M., Hiramatsu,K. and Kikuchi,H.
            Direct Submission
TITLE       Submitted (30-JAN-2001) Director-General, Biotechnology Center,
            National Institute of Technology, Shibuya-Ku, Tokyo 151-0066, Japan
            (E-mail:bio@nitech.go.jp, URL:http://www.bio.nitech.go.jp/,

```

```

COMMENT      Tel:81-3-3481-1933, Fax:81-3-3481-8424)
FEATURES
    source
        On Jun 12, 2001 this sequence version replaced gi:13701256.
        Location/Qualifiers
            1..301550
                /organism="Staphylococcus aureus subsp. aureus N315"
                /mol_type="genomic DNA"
                /strain="N315"
                /sub_species="aureus N315"
                /db_xref="taxon:158879"
                /gene="SA1293"
                /complement(1249..947)
                /gene="SA1293"
                /note="ORFID:SA1293"
                /codon_start=1
                /transl_table=11
                /product="conserved hypothetical protein"
                /protein_id="BAB42553.1"
                /db_xref="GI:13701256"
                /translation="MSIIFMTIYFLVLPPLMAOHKYSNTERYSQVRSCTGR
                EVALLEHANGIYDEVVKGFLDHDPRKKVSLSPANDRPSVAGTAIAHEVG
                HAIDHQGYWELFRFAALVPANGLSSLSITIMILITLIGSAGSTALMIGALM
                SLAVLFSTVLPVEFDASSRAMKOITALNTVNEKEYHAKKVLSSAAATYVATAVAV
                AELVRIILARSDD"
                /complement(1001..1588)
                /gene="SA1294"
                /complement(1001..1588)
                /gene="SA1294"
                /note="ORFID:SA1294"
                /codon_start=1
                /transl_table=11
                /product="conserved hypothetical protein"
                /protein_id="BAB42554.1"
                /db_xref="GI:13701260"
                /translation="MTINFQWYLYLQYRSYMLMLICNLIGMTIYGYWGLSHTPM
                HFKIYDPSPTALILYVLSISLILQKNSIIFALAVFLFKGIWAVLMTFLIEG
                GDITNGVLVMSHSIMAVOALFYPPFRKSMIGISVAMIMWFLNDYIDYFILQPPY
                DRTTHVMQIGVLCSSVFCLLIYELKRLCK"
                /complement(1578..2153)
                /gene="SA1295"
                /complement(1578..2153)
                /gene="SA1295"
                /note="ORFID:SA1295"
                /codon_start=1
                /transl_table=11
                /product="conserved hypothetical protein"
                /protein_id="BAB42555.1"
                /db_xref="GI:13701261"
                /translation="MSETLNQIKESFIEYLLFQYRKRSLAVWLVYIKVNEAKIAN
                HVDPRKIHMHETLEAPVGSNASHAIOFTKRNITLNTNPNTPIYIANHNCAEDIOHFA
                NYSKREORLDILYNOLESFSTOTTLADNSMAIDRRHALLDILDLNRDLSLQNM
                EKORFYOLTLNLTAKVNHNFEDLADD"
                /complement(2167..3411)
                /gene="SA1296"
                /complement(2167..3411)
                /gene="SA1296"
                /note="ORFID:SA1296"
                /codon_start=1
                /transl_table=11
                /product="conserved hypothetical protein"
                /protein_id="BAB42556.1"
                /db_xref="GI:13701262"
                /translation="MEDYKILIDINLQKLENIDSNVNEAITTDNDALFLIGETLVN
                FELMPQGLEVRVLYHKYPODESLIYFIEGLMSENGQDEALEYLSVSPERKMLEE
                ADLYOQINNMVAIDKLDQALELPNDPIIHFALEMLYDQGLRATSEVLEVLGG
                EYQVNGVNLFSRMADCSQSGNSDAIRLVDINEDEMTSEDYLRKASYNKNDTGE
                AIKIMTLLSDPDYIOGYLYLQSLYENENYPDALETGKBEGLRSQFKELMTWTCG
                LEIHGDANSGVLLKQALEVDNAYGPELILSDYLARNEDYEAIIELLTGYVEEDLD
                PTFWHLAFAYGGEERDKEAHPFELAYPTMKTNIDPMSDYFYFLIEIGQEQDAITLL
                NQLELSPSNMNMDELSPLOS"
                /complement(3418..4716)
                /gene="aica"
                /complement(3418..4716)
CDS
gene
CDS

```


TITLE	Genome and virulence determinants of high virulence
JOURNAL	Community-acquired MRSA
MEDLINE	Lancet 359 (9320), 1819-1827 (2002)
PUBMED	22040717
REFERENCE	2 (bases 1 to 333750)
AUTHORS	Director-General, Biotechnology Center, Aoki, K., Oguchi, A., Nagai, Y., Asano, K., Iwama, N., Baba, T., Kuroda, M., Hiramatsu, K. and Kikuchi, H.
TITLE	Submitted (06-MAR-2003) Director-General, Biotechnology Center, National Institute of Technology and Evaluation, Biotechnology Center, 2Chome 49-10 Nishihara, Shibuya-Ku, Tokyo 151-0066, Japan (E-mail: bio@nite.go.jp, URL: http://www.bio.nite.go.jp/, Tel: 81-3-3481-1933, Fax: 81-3-3481-8424)
JOURNAL	Location/Qualifiers
FEATURES	1. 333750
source	/organism="Staphylococcus aureus subsp. aureus MW2"
	/mol_type="genomic DNA"
	/strain="MW2"
gene	/db_xref="taxon:196620"
	complement(278. .619)
CDS	/gene="MW1341"
	complement(278. .619)
	/gene="MW1341"
	/note="ORFID: MW1341"
	/codon_start=1
	/transl_table=1
	/product="hypothetical protein"
	/protein_id="BAB95206.1"
	/db_xref="GI:21204510"
	/translation="MTRITEVFDNDEQEDIVPAKKTKSGSEAPRIIDYHSHIIOE EKQVNDIGNIDFDKLNQIFVPMNFKETIDIEKTHFGICRMKTFKTELINMA SYQRLNKKQV"
gene	complement(624. .1283)
CDS	/gene="nth"
	complement(624. .1283)
	/gene="nth"
	/note="ORFID: MW1342"
	/codon_start=1
	/transl_table=1
	/product="endonuclease-like protein"
	/protein_id="BAB95207.1"
	/db_xref="GI:21204511"
	/translation="MWSKKALEMIDIVANMPDAECLEKHNPELTIAVLISAQCT DVLNRYVTELEFKYKTPEDYLVASDEELMNDIRSIGLRNKKIRKLCQSLDOFN GEIPQTRKELESAGVGKRTKATVAVSVADDEPSLAVDTHVERVSRGICINRMKNVRO VEDRLCSVIPRDMNNSHOLIFFGRHCLARRPKDICPLLEDCREGQKKASLKE A"
gene	complement(1273. .1959)
CDS	/gene="MW1343"
	complement(1273. .1959)
	/gene="MW1343"
	/note="ORFID: MW1343"
	hypothetical protein, similar to chromosome replication initiation protein dnaB"
	/codon_start=1
	/transl_table=1
	/protein_id="BAB95208.1"
	/db_xref="GI:21204512"
	/translation="MDKYOLKARPVIRRELDHYSDGLDEODLVILKLIVASETS NKOPSTELKSGSTMOBDITWVIONLORLELLOKDEGRTEYMNIDPEPEKL SHLKQOSMETKQNSKEKFKOLFVLETPARPSPREIETLNMIDVDHDSAIIO AALDEANSLKLSFKITMDRITLLMKNNKNTTIDSKIREKFNKKMHTVKTVPKD WLNGENLDGK"
gene	complement(2287. .3579)
CDS	/gene="asn5"
	complement(2287. .3579)
	/gene="asn5"
	/note="ORFID: MW1344"
	/codon_start=1
	/transl_table=1
	/product="asparaginyl-tRNA synthetase"
	/protein_id="BAB95209.1"
	/db_xref="GI:21204513"
	/translation="MKTITIKQAKDHLNDVTTIGAMLRNRSCKIAPILQLRDGTGFMQ GVVKSVEDEEVEFLAKEITQESLVTYGTITEDNSDGLGYEMQVKSIEVSEADNP ITPKNHGTPELDHRLMLRSKKQAAVAKIRNEVIAIATEYEPFNKQGFTRVPLTAS APEGTSLEFHTKYFDQDAFLSOSGLYLEAANAAGKVPFGPTPRASKSTRHLLIE FVMIGEMAFTHNAESLEIOQVYTHVKSVLNCLELTKLIERDTSLEKVPYAPFPR ISYDAIEFLKSGEFDIEMGEDGADHETIAINHHDLVPTFTYTRPKIKFQYQPNP EHEETVLCADLIAPBGYEITGSEERDDELDDPKREHGEDBEAAYIYDLNRYS VPHGCFGLEERTVAMISGVNHEFAPFRLNRLKP"
gene	complement(3901. .6594)
CDS	/gene="dmg"
	complement(3901. .6594)
	/gene="dmg"
	/note="ORFID: MW1345"
	/codon_start=1
	/transl_table=1
	/product="probable ATP-dependent DNA helicase dmg"
	/protein_id="BAB95210.1"
	/db_xref="GI:21204514"
	/translation="MGMAITVAVDLETGNGQDFDITDITGTPYRNQITDYHSMI RTNLEIPPEIQLTSTENMLQAPTFNVAQDITDKINDCTLVANNDPLNTKKA FKDNIQYRPKRVITDLEIFRIAPPTDKSYQSELEAGITLVANRADDAATYAK LMILAEKFEKPLDITLKOYLKQDKDYDIFEVNRODARPLDKSYKEEOII YRKQYDFKPTTNGSLSKSLYSAVDQDLTYRPOQLAETIIDLQMLSEKAMIEA SIGSGKSLAYLLAALNIEGKHVMISTRTLRQSLAETIIPAMNALLNFKXIALI IKSKSDYISGLISQILKDDTSNVEVILMOLLITETPEGDIQELNLKQGRWF DKIETVYPARHVDVYNFPIKRNANONIGITTHMALHSDVENSIIYOLFDCIYDVA HRLPVALNOVTNELSYADI KYOLGLKQNMENKILKAIIDOLEKRIELKIDIAIDI FGKASMEIHEIMEDLPSTITITINDSOVDYDDHRRPNVPEPTKQILDAIIDI KLKRTLEIFNCSISKRTVRSRKQOLYLKQKFRKIDQSLKAGTSISIKNSQKSTR LTVKDIAYADVLITKQVLEKRSLSLTISGLTRKHSFEAKQFNQDVFNTFVNTSL QSAKTSVFITPSDVAISTYKRIDVYASIVSITETITITSKCLVLETRNMMHWQ DMLNLFPEDEDVVLTOOONONKRIQOEVFNPAKRALIGLSTSFEGSFOANGIKCYA IAKLFPNKNHAKYMLADSESTSEFEVYLPAAVYRFGQLRLRLRSBENDGIIYSFD DRLINSNYKNFEQTLNRYRQKKDIDQEGKLLRQYOKKKK"
gene	complement(6618. .7589)
CDS	/gene="MW1346"
	complement(6618. .7589)
	/gene="MW1346"
	/note="ORFID: MW1346"
	/codon_start=1
	/transl_table=1
	/product="truncated hypothetical protein"
	/protein_id="BAB95211.1"
	/db_xref="GI:21204515"
	/translation="MSKYSDVAVQLLYKKNPNTISGOSIAESLNSRTAVKAYIDQK LEGKIDSVNHHGHLQQLPDYIOGIIQDYKSSALFQFSVYDSIDSTOLAAKSL VGNOSFFILSDEOTKGRNFRNRMSSSGGGLMNSVYLPRVAFMSIKFNPLIAG IRDAIOHFSODEEVKVPNDIYIDNCKVCGFLTEVANNNDGIEAIIIGIGINLQOE NPDESIRHRATSVOLHDKNKLDRYOFLEBLIEIKRKNOPITLPFSFIREYIAASN IWNRTILFTENDQFKQCAIDLDYDGLIVREBAGESHRLISAQIDP"
gene	complement(7576. .8778)
CDS	/gene="MW1347"
	complement(7576. .8778)
	/gene="MW1347"
	/note="ORFID: MW1347"
	hypothetical protein, similar to poly(A) polymerase"
	/codon_start=1
	/transl_table=1
	/protein_id="BAB95212.1"
	/db_xref="GI:21204516"
	/translation="MDSLSEBQARPLBEDIONGCEATVYCGSVADYIMGNINHDID TTSATPDESISHTIPVCKREHCTTNVFNDENEVTFTRNEEDYDHRPSCVTV RDLYEDLRQROFTNMAIAMOTAYKRLDYDGGQDIIENNIITRVGIAERFQEDLRLNI RCILRFQSLSFDAITETFEAMRIOMADIFSLIRIVLETLKLNKINVEKSPHMLS LKAFNMPYEFELDMQINVTBAIDELIAIVSKFINSLRPLKLSNQVNDIINO YIOINNALPISITTEQOLKMFVYDQTHLIRNVAADVIKANDIOGHEPLLVNQTID ETLRLRPLHNKDKMNVGCVLMAHLNAKSGMLKDVLIQDIEAIVTGVNSEETELK MVDNHVKI"
gene	complement(8783. .9925)
CDS	/gene="MW1348"
	complement(8783. .9925)


```

/gene="MW1348"
/Note="ORFID:MW1348
hypothetical protein, similar to lipopolysaccharide
biosynthesis-related pr homolog"
/codon_start=1
/transl_table=11
/protein_id="BAB95213.1"
/db_xref="GI:21204517.1"
/translacion="MKIGTICPSMGSGIINTELCIKIAERHEVHTSNIPRI
KPLENMFHOVNOYAVFOYPPYDITLSTKIAEVIKEGDLIMHAYPAHICIL
AREMSKMDIKMTLHGTDTVGLDHSOGAKIKEIKSDIVTSVSKSLAQETHET
ETNKEIIPINPARENFPKNTALKSOFGLAPDEKVIHVSINROVKRDI1IEF
AKVREKIPSLILLDGPELPMKQLTKEINAEVDFGLKQDCVSEFYQSDVLELL
SEKESFGLTLEAMKTVGVIGSNAGIKKEIKHGEFGFVVDGCDSDAYAIRLE
DKVYNLKONMLADIAERGSSELIIDQREYVYOKKLNHNKSKE"
/complement(10167..10484)
/gene="MW1348"
/complement(10167..10484)
/gene="MW1349"
/Note="ORFID:MW1349"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="BAB95214.1"
/db_xref="GI:21204518"
/translacion="MKSVMOREVDEYIGOFKGYFSPLANLARLDEVEGLAREIN
HTYERKKKSEANNTIKAEIDNLFVCLCLANSNGIDMTESPNETMEKFNTRDKNR
ERK"
/complement(10820..11518)
/gene="MW1350"
/complement(10820..11518)
/gene="MW1350"
/Note="ORFID:MW1350"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="BAB95215.1"
/db_xref="GI:21204519"
/translacion="MSITFMITYEYILNVLPMQHKKSNEYKRSOVRSTSGCTGR
EVALEILHANGITDVEYVKGESFLTDHOPKRVLSSPANDRPSVAGTAAHEVG
HAIDHOGYFLRRALVPVANNLSSLSMIIMIGLITLITAGISAFGSTALMIGGLM"
Query Match 100.0%; Score 387; DB 1; Length 333750;
Best Local Similarity 100.0%; Pred. No. 1.2e-77;
Matches 387; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

RESULT 4
AP003362
LOCUS
DEFINITION
Staphylococcus aureus subsp. aureus Mu50 genomic DNA, complete
sequence, section 5/9.
ACCESSION
AP003362.2 GI:14247083
VERSION
AP003362.2
KEYWORDS
ORGANISM
Staphylococcus aureus subsp. aureus Mu50
Bacteria; Firmicutes; Bacillales; Staphylococcus.
REFERENCE
1
Kuroda,M., Ohta,T., Uchiyama,I., Baba,T., Yuzawa,H., Kobayashi,I.,
Cui,L., Oguchi,A., Aoki,K., Nagai,Y., Iino,J., Ito,T., Kanamori,M.,
Matsumaru,H., Maruyama,A., Murakami,H., Hosoyama,A.,
Mizutani-U.Y., Takahashi,N.K., Sawano,T., Inoue,R., Kato,C.,
Sekimizu,K., Hirakawa,H., Kuhara,S., Goto,S., Yabuzaki,J.,
Kamihisa,M., Yamashita,A., Oshima,K., Furuya,K., Yoshino,C.,
Shiba,T., Hattori,M., Ogasawara,N., Hayashi,H. and Hiramatsu,K.
Whole genome sequencing of methicillin-resistant Staphylococcus
aureus
Lancet 357 (9264), 1225-1240 (2001)
JOURNAL
MEDLINE
PUBMED
21311952
11418146
REFERENCE
2 (bases 1 to 346900)
Ohta,T.
Direct Submision
Submitted (28-Feb-2001) Toshiko Ohta, University of Tsukuba College
of Medical Technology and Nursing, Department of Medical
Technology, 1-1-1 Ten-noodai, Tsukuba, Ibaraki 305-8577, Japan
(E-mail:tohta@tsukuba.ac.jp, Tel:81-298-53-3454,
Fax:81-298-53-3454)
On May 29, 2001 this sequence version replaced gi:13875626.
COMMENT
Location/Qualifiers
1..346900
/organism="Staphylococcus aureus subsp. aureus Mu50"
/mol_type="genomic DNA"
/strain="Mu50"
/sub_species="aureus Mu50"
/db_xref="taxon:158878"
12..263
/gene="SAV1313"
12..263
/gene="SAV1313"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="BAB57475.1"
/db_xref="GI:14247084"
/translacion="MLTLIKAEDEOVIIIEYIPEDISNGKCSYTFNNKDAEVIDFS
LSEIEEEVFLMYRNKSFVVRDEIEROEPENKIAMY"
987..1181
/gene="SAV1314"
987..1181
/gene="SAV1314"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="BAB57476.1"
/db_xref="GI:14247085"
/translacion="MNLOSXYIGKLVKLTNNKILGKVIDFDKVDNFDGYSIEID
FGRIYDISENKIKIVQK"
complement(1725..2750)
/gene="SAV1315"
complement(1725..2750)
/gene="SAV1315"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="BAB57477.1"
/db_xref="GI:14247086"

```

/translation-"MISFENDYLGSAHEKYLKRLVDPTNLVQASGCGFDOPTAOAIETI
KOTIDCPNATIRFLVGGTQTNQVAVINSMLSEYEGVISADTGVAHVAGALISYSHKV
ITIPSEKRVASADVETMETPKSDFKRDMHVPFGVYISHPTXGTLYSKSELEELC
KVCQYOLPLPMDGARLGYGLMSQSDMTIKDIKACDVFYIGGTIGALGCEALVFT
KNNEKOPTTRIKKHGALILAKGRLTGTOFLEPTDNLVFNISRHAEAMNKKDKFKN
KGYRLVPSPTNOOFFILSNKEIKAELOKVPFVMEKVDQDHVVFATSNMTTEENL
NKLEELI"
gene 3020..3217
/gene-"SAV1316"
CDS 3020..3217
/gene-"SAV1316"
/codon_start-1
/transl_table-1
/product-"hypothetical protein"
/protein_id-"BAB57478.1"
/db_xref-"GI:14247087"
/translation-"MEQIKIKTFYAEITELLEKKINNAFLSSEATNLKLVNITKEIE
BRTFPNNEERNALITLSVKN"
gene 3238..4755
/gene-"SAV1317"
CDS 3238..4755
/gene-"SAV1317"
/codon_start-1
/transl_table-1
/product-"cardiolipin synthetase homolog"
/protein_id-"BAB57479.1"
/db_xref-"GI:14247088"
/translation-"MFCVPPFPGIKETMRTPFNSDLGTFTIILAGFIINLVAFIIT
FLERNRPASTMAVLVPLVPLIGLILFPGRTVSARKLNKNGNAGTLPDCLIK
QOIESFDGNGTQNKQYOKRHDVLSMLMDQGFLENNKVDHPTDGDNDLDOVLDK
IKNAKETIHEITYTFALDGSKRLHLAEKLGKQLEVKILYDVGSKVNAKNDHF
KSLGGEVAFVASKLPLNFMNNRNRKRIIVDQGLGVGGTIDETELGJGKRLGY
KDTLRIIGDADVALQKFLIDMNSQAHROPQFEDVQYKPKKNGPLGNSP101ASGP
ASDMQIIEGYGATKIMSAKSVYLSQSPYIPDINSYINAIKIAKSCVDVHLAIPEKPD
HPLVYMAFYSASDLISGKVIYTYENGFTHSKMCCLDEIVASVGTANDFSPFELNF
EVNAFVYDENLAKDLRVAEHDITKSKQLTESYANRPLSVKFEKESLARLVSPIL"
gene 4939..5436
/gene-"SAV1318"
CDS 4939..5436
/gene-"SAV1318"
/note-"ATP-binding protein"
/codon_start-1
/transl_table-1
/product-"ABC transporter homolog"
/protein_id-"BAB57480.1"
/db_xref-"GI:14247089"
/translation-"MIOISINIKSFNRRCVLYKNISPDIEGKCIALIGKNGAKSTLI
DLIGNVANSGEIFDKDKLISNSRSIMQFMFPDQKVIETIINLYGSPFENPLPL
BEIILTKFSSQLNQFVNKLSCGQRLDPLVLSLIGPQLILDEPTMDIEIREY
FMSIY"
gene 5461..5838
/gene-"SAV1319"
CDS 5461..5838
/gene-"SAV1319"
/note-"ATP-binding protein"
/codon_start-1
/transl_table-1
/product-"ABC transporter homolog"
/protein_id-"BAB57481.1"
/db_xref-"GI:14247090"
/translation-"MLTYSHIIEVEYRMSDKILIEGCEIILNDSTSHRTNDOOIST
LSDETRILKDKDLVIOKNHNGTIIITISVNDTILVLOQLHNLDDIEIQRKSVISV
DSYFNKKRQGSNDYTKLENRI"
gene 5807..6538
/gene-"SAV1320"
CDS 5807..6538
/gene-"SAV1320"
/note-"similar to ABC transporter integral membrane
protein"
/codon_start-1
/transl_table-1
/product-"hypothetical protein"
/protein_id-"BAB57482.1"

/db_xref-"GI:14247091"
/translation-"HILSYLIEREIVMRKKTLLIILSPVITFYLTFTSLLEPEDV
KMPFYKEVYWTSTVLSLFTSPDIDINEKNENRQMLAMTPTPTFSYISNVK
TKMGEFIALVIFVNGHFRYKGVANSAVOMLESGFELMLGSLITGLIFSLNDIOK
TSLMALNIVIGLAVLGGMLPFIINFPNMLDHAVHVPYSYLRKLGVDIASNNHNLIS
FALLIYVIGSIIAVYCSHFRRAE"
gene 6542..7633
/gene-"SAV1321"
CDS 6542..7633
/gene-"SAV1321"
/note-"similar to two-component sensor histidine kinase"
/codon_start-1
/transl_table-1
/product-"hypothetical protein"
/protein_id-"BAB57483.1"
/db_xref-"GI:14247092"
/translation-"MKPLKDTISAEISILYLPIPIAGIPFENEYCPKMLYIISVIF
SISYLIVVNNRLNTLMFYILLIHVFIICYFVSVHPLSLFFPYSAVAPFEN
NYKRTAPNFILTMITICITITVILYNNFVAMVYVVISLIMDPKMKRREYKE
IARKNRHNTLADDERHRICODLHRTGHVFSASLSKSLAKLIDTPEVEYKAEIL
AIKLSRESLAKVREIYDVKLPSTFIEIDIRKVLADADIDTFPKKELAOVLSPTK
OSMLVMTREALINNVIKHANASKVHGKLTVNNHKLMLIEDGKIDSDCEVKSISO
RVQHLNGTCLAVDSTNGTKRIIEISTGIA"
gene 7630..8232
/gene-"SAV1322"
CDS 7630..8232
/gene-"SAV1322"
/note-"similar to two-component response regulator"
/codon_start-1
/transl_table-1
/product-"hypothetical protein"
/protein_id-"BAB57484.1"
/db_xref-"GI:14247093"
/translation-"MISLIADENMLRQAAVOLIKIHGOFEILLADTDNGLAKLIE
EYNPVNIIDIDMPGTGLEVAELERKHLIRVYIYTPKRGYIEKAAVNDVDAVY
LKRSTSEELVETINKNGEKESATLMTSEFYDKNLTPKDOIIVLREIGNLSKEI
SEKLPLTDGTVENYTSVYIIDKLFADNREFDAWKANKENGW"
gene complement(8345..8533)
/gene-"SAV1323"
CDS complement(8345..8533)
/gene-"SAV1323"
/codon_start-1
/transl_table-1
/product-"hypothetical protein"
/protein_id-"BAB57485.1"
/db_xref-"GI:14247094"
/translation-"MNSNENVTLLAVIINIIIGIFLQJFFRLPLVTAVALSILLGI
FVGFIYLVISFYNNRKN"
gene 8672..9205
/gene-"nuc"
CDS 8672..9205
/gene-"nuc"
/note-"SAV1324"
/codon_start-1
/transl_table-1
/product-"thermonuclease"
/protein_id-"BAB57486.1"
/db_xref-"GI:14247095"

Query Match 100.0% Score 387 DB 1: Length 346900:
Best Local Similarity 100.0% Pred. No. 1,2e-77;
Matches 387; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GATCTTCTTCTCTTCCACCAAAATGAGAAACAATGCTATACAGTACCAAGACCTA 60
|||||
DB 200303 GATCTTCTTCTCTTCCACCAAAATGAGAAACAATGCTATACAGTACCAAGACCTA 200362
|||||
QY 61 AACCATGTGACCTGTAATCGGATACGGTACCAACCAATCGTAATGATACAAATCATPAC 120
|||||
DB 200363 AACCATGTGACCTGTAATCGGATACGGTACCAACCAATCGTAATGATACAAATCATPAC 200422
|||||
QY 121 CGTCTGACCAATTCATATATTACTTGTGTAACCGCTAATACGACCGGTTTTTAA 180
|||||

```
Db 200423 CGTCGTACCGATTCCATTTATCTACTTTGTAAACCGCTAATACGACGGTTTTTTAG 200482
OY 181 ATTTGTATTAATAATTTGAGCAGACATTTCAATCGCTTGTGTCAATCGCTTCAGCAGTTAA 240
Db 200483 ATTTGTATTAATAATTTGAGCAGACATTTCAATCGCTTGTGTCAATCGCTTCAGCAGTTAA 200542
OY 241 CCATTAATAATAATTAACATCCGCTTCATCTATGCGGATTTCTGCGTGCCTCTAATTTGTG 300
Db 200543 CCATTAATAATAATTAACATCCGCTTCATCTATGCGGATTTCTGCGTGCCTCTAATTTGTG 200602
OY 301 TTGTGAATGTCATCACCATAATTCATACACCTGTATCAATATTAATTAATCATGTG 360
Db 200603 TTGTGAATGTCATCACCATAATTCATACACCTGTATCAATATTAATTAATCATGTG 200662
OY 361 TTACCATTCACCGAAGATTAATAC 387
Db 200663 TTACCATTCACCGAAGATTAATAC 200669

RESULT 5
AF270032 3269 bp DNA linear BCT 01-AUG-2000
LOCUS Staphylococcus epidermidis strain SRI clone step.1042f07 genomic
DEFINITION sequence.
ACCESSION AF270032
VERSION AF270032.1 GI:9623936
KEYWORDS Staphylococcus epidermidis
SOURCE Staphylococcus epidermidis
ORGANISM Bacteria; Firmicutes; Bacillales; Staphylococcus.
REFERENCE 1 (bases 1 to 3269)
AUTHORS Kimmerly,W.J., Taylor,J.David., Nelsen,A.J., Godlewski,M.M.,
Rubino,M.A., Nelsen,F.J., Rivers,P.R., Torrella-Miller,I.,
Listenbee,S., Ashanti,C., Altschuller,G., Mammo,L., Shepherd,N.S.,
Fuchs,R., Fleming,T., Guan,X., Du,L., Cain,D.H., Miller,G.S. and
Furdon,P.J.
Transposon-mediated sequencing of the staphylococcus epidermidis
genome
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 3269)
AUTHORS Taylor,J.David., Kimmerly,W.J., Nelsen,A.J., Godlewski,M.M.,
Rubino,M.A., Nelsen,F.J., Rivers,P.R., Torrella-Miller,I.,
Listenbee,S., Ashanti,C., Altschuller,G., Mammo,L., Shepherd,N.S.,
Fuchs,R., Fleming,T., Guan,X., Du,L., Cain,D.H., Miller,G.S. and
Furdon,P.J.
Direct Submission
JOURNAL Submitted (22-MAY-2000) Departments of Genomic Sciences and
Bioinformatics, Genetics Directorate, Glaxo Wellcome, Inc., 5 Moore
Drive, Research Triangle Park, North Carolina 27709-3398, USA
FEATURES
SOURCE 1. 3269
/organism="Staphylococcus epidermidis"
/mol_type="genomic DNA"
/strain="SRI"
/db_xref="taxon:1282"
/clone="step.1042f07"
BASE COUNT 986 a 616 c 451 g 1216 t
ORIGIN
Query Match 67.8% Score 262.2; DB 1; Length 3269;
Best Local Similarity 79.8% Pred. No. 3.1e-49;
Matches 309; Conservative 0; Mismatches 78; Indels 0; Gaps 0;
OY 1 GATCTTCTCTCTTCCACCAATAGAGAAACATGTCATCAACAGTACCAACCTA 60
Db 2297 GATCTTCTCTCTTCTTAAAGTTTCAACAACATGTCATCAACAGTACCAACCTA 2356
OY 61 AACATGTGACCCGATATCGATAGCGTTTCACCAATCTTAATGAATGAAATCATCA 120
Db 2357 ATCCATGTGACCCGATATCGATAGCGTTTCACCAATCTTAATGAATGAAATCATCA 2416
OY 121 CGTCGTACCGATTCCATTTATCTACTTTGTAAACCGCTAATACGACGGTTTTTTAG 180
Db 121 CGTCGTACCGATTCCATTTATCTACTTTGTAAACCGCTAATACGACGGTTTTTTAG 180

Db 2417 TATCATTTACGATTTCAAGATTAATCAACTTTATTCACAGCTAATACAGAGTTTCTAG 2476
OY 181 ATTTGTATTAATAATTTGAGCAGACATTTTCATCGCTTGTGTCAATCGCTTCAGCAGTTAA 240
Db 2477 ATTTGTATTAATAATTTGAGCAGACATTTTCATCGCTTGTGTCAATCGCTTCAGCAGTTAA 2536
OY 241 CCATTAATAATAATTAACATCCGCTTCATCTATGCGGATTTCTGCGTGCCTCTAATTTGTG 300
Db 2537 CCATTAATAATAATTAACATCCGCTTCATCTATGCGGATTTCTGCGTGCCTCTAATTTGTG 2596
OY 301 TTGTGAATGTCATCACCATAATTCATACACCTGTATCAATATTAATTAATCATGTG 360
Db 2597 TTGTGAATGTCATCACCATAATTCATACACCTGTATCAATATTAATTAATCATGTG 2656
OY 361 TTACCATTCACCGAAGATTAATAC 387
Db 2657 TTACCATTCACCGAAGATTAATAC 2683

RESULT 6
AX145350 3269 bp DNA linear PAT 31-MAY-2001
LOCUS Sequence 4072 from Patent WO0134809.
DEFINITION AX145350
ACCESSION AX145350
VERSION AX145350.1 GI:14283915
KEYWORDS synthetic construct
SOURCE synthetic construct
ORGANISM artificial sequences.
REFERENCE 1
AUTHORS Kimmerly,W.J.
TITLE Staphylococcus epidermidis nucleic acids and proteins
JOURNAL Patent: WO 0134809-A 4072 17-MAY-2001;
GLAXO GROUP LIMITED (GB)
FEATURES
SOURCE 1. 3269
/organism="synthetic construct"
/mol_type="genomic DNA"
/db_xref="taxon:32630"
/note="synthetic nucleic acid sequence"
BASE COUNT 986 a 616 c 451 g 1216 t
ORIGIN
Query Match 67.8% Score 262.2; DB 6; Length 3269;
Best Local Similarity 79.8% Pred. No. 3.1e-49;
Matches 309; Conservative 0; Mismatches 78; Indels 0; Gaps 0;
OY 1 GATCTTCTCTCTTCCACCAATAGAGAAACATGTCATCAACAGTACCAACCTA 60
Db 2297 GATCTTCTCTCTTCTTAAAGTTTCAACAACATGTCATCAACAGTACCAACCTA 2356
OY 61 AACATGTGACCCGATATCGATAGCGTTTCACCAATCTTAATGAATGAAATCATCA 120
Db 2357 ATCCATGTGACCCGATATCGATAGCGTTTCACCAATCTTAATGAATGAAATCATCA 2416
OY 121 CGTCGTACCGATTCCATTTATCTACTTTGTAAACCGCTAATACGACGGTTTTTTAG 180
Db 2417 TATCATTTACGATTTCAAGATTAATCAACTTTATTCACAGCTAATACAGAGTTTCTAG 2476
OY 181 ATTTGTATTAATAATTTGAGCAGACATTTTCATCGCTTGTGTCAATCGCTTCAGCAGTTAA 240
Db 2477 ATTTGTATTAATAATTTGAGCAGACATTTTCATCGCTTGTGTCAATCGCTTCAGCAGTTAA 2536
OY 241 CCATTAATAATAATTAACATCCGCTTCATCTATGCGGATTTCTGCGTGCCTCTAATTTGTG 300
Db 2537 CCATTAATAATAATTAACATCCGCTTCATCTATGCGGATTTCTGCGTGCCTCTAATTTGTG 2596
OY 301 TTGTGAATGTCATCACCATAATTCATACACCTGTATCAATATTAATTAATCATGTG 360
Db 2597 TTGTGAATGTCATCACCATAATTCATACACCTGTATCAATATTAATTAATCATGTG 2656
OY 361 TTACCATTCACCGAAGATTAATAC 387
Db 2657 TTACCATTCACCGAAGATTAATAC 2683
```

db 2657 TTAACCACTCCGACATGATTAATATAC 2683

RESULT 7
AE016747 300892 bp DNA linear BCF 01-JAN-2003
LOCUS Staphylococcus epidermidis ATCC 12228, section 4 of 9 of the
DEFINITION complete genome.
ACCESSION AE015929
VERSION AE016747.1 GI:27315369
KEYWORDS
SOURCE Staphylococcus epidermidis ATCC 12228
ORGANISM Staphylococcus epidermidis ATCC 12228
REFERENCE Bacteria: Firmicutes; Bacillales; Staphylococcus.
1 (bases 1 to 300892)
AUTHORS Zhang, Y., Ren, S., Li, H., Fu, G., Lu, L., Lu, G., Jia, J., Tu, Y.,
Qin, Z., Chen, Z., and Wen, Y.
TITLE Direct Submission
JOURNAL Submitted (05-NOV-2002) Chinese National Human Genome Center at
Shanghai, 250 Bi Bo Road, Shanghai 201203, China

FEATURES
source location/Qualifiers
1..300892
/organism="Staphylococcus epidermidis ATCC 12228"
/mol_type="genomic DNA"
/strain="ATCC 12228"
/db_xref="ATCC:12228"
/db_xref="taxon:176280"
113..346
/locus_tag="SE0907"
113..346
/locus_tag="SE0907"
/codon_start=1
/transl_table=1
/product="acyl carrier protein"
/protein_id="AA04504.1"
/db_xref="GI:27315370"
/translation="MENEDKVKDIIVRLGVADKVTEDAFKFDGLDGLDIELVH
ELEDFGTEIPDESEAKINTVGAVKYINSLER"
461..1198
/locus_tag="SE0908"
461..1198
/locus_tag="SE0908"
/codon_start=1
/transl_table=1
/product="Rnae III"
/protein_id="AA04505.1"
/db_xref="GI:27315371"
/translation="MANQKKKEVHNFOKFTDKMKSGLHFKFNIDLYQOAFSSHSFI
NDFNNRLNHNRLLEFLGDVAVLELTVSRYLFDNRPHLPBGNTLKHRTATVCEPSLYIF
ANKIKNELILLGKEERTGRTAPSLISDAFEAFGALYLDQGLDSVTFPAEXYIF
YVEDELGVVDFPTFOEYVHSQNGDVTYQLIKREGRPHHRLFTSEVILEKNAVAE
GKGTKKSESOAKAQAQVKKMKKNSL"
1325..4894
/locus_tag="SE0909"
1325..4894
/locus_tag="SE0909"
/codon_start=1
/transl_table=1
/product="chromosome segregation SMC protein"
/protein_id="AA04506.1"
/db_xref="GI:27315372"
/translation="MVYLKSTIDAGFSPADHNNVOPDKGVTAIVGPNSSGKSNITDA
TKVVLGDSASLSGAKKEDIIPSGAHRRAQNFAPAYKLLDNHSOKLIDSEELVVT
RLVLRSGSEITLNDNRRLKDIIDPLDGLGCKEAFSIIISGRVDEILNAPIRRKQ
LIESAGVATKKRKRASISQILDTHTENLRNVDIILYDEGRPELKEEAATAKYKQ
ROQNDYDEKLVNLELVKATENVPOLSDHTEDNRDLNRLNLSQAQKEGQAOIQLQAKYKQ
SIKNEKAQNEKLALNKKQKLNKEVQELSLYLSDEQHDKELEIKNSYTLASE
OSDVNNDIRPLEHTINENKASRLDSRLVEARNOAKDIOQNTOTOKEYSSKSMK
KVEQNIQOLEQOFLTSKRLSEYENKLYQAYRNEKLRSIDSLATQEDDYTFPFGV
KHILKAKDELRGTHGVAEVIYVPSMTQAITTAAGASLQHVYIYNEMKQDAIOYL
KORGLGRATFLPLNVITOPRHVADIKVAAGSGFTINISDAININSAKTONITENLIG
MTIYENLKHANELARIRTRIRLVLTGDEVNPGSGMTGGGAKRTKSTLSOKDELST

gene
CDS
4891..6117
/locus_tag="SE0910"
4891..6117
/locus_tag="SE0910"
/codon_start=1
/transl_table=1
/product="signal recognition particle"
/protein_id="AA04507.1"
/db_xref="GI:27315373"
/translation="MSFRKIKDRFSKSEPDIOKIDFTVDSNANSDDSNPNDS
EYKPRKKRKSLSEADFDGCLISTEDFEIEAKITAKTKAGLEKSRONFDQNLN
IATRKVDEDFEALREMLITADGFNTVMKLTDELTEKQRRNIDETEDLREVIYK
IVETIYQEDDSSEANNIEGKRLNVLNMGVGVKRTTIGKLVRYQOGBKVMYLAAG
DNFRACAIQQLNMGVEBVGVEVVSQESDPAVAYDAINAAKSGVDILICTAGRL
QKNSNLMOELDKMKRYINRAIPDAPHEALLDATTQNALSQARSKFEVTNAGVL
TKLDGTAKGIVLAIINELHIYVKKYVIGGEIMDOLQFPNPEYVYGLFADMIQNE
PEIISRNSVESEEGN"
6119..6451
/locus_tag="SE0911"
6119..6451
/locus_tag="SE0911"
/codon_start=1
/transl_table=1
/product="conserved hypothetical protein"
/protein_id="AA04508.1"
/db_xref="GI:27315374"
/translation="MGNDLVATLMMNTLFPYOSLLTNKKNYLEFFYLQDYSSEI
ADFEVROAVDNIRRTGDLVEDYESLRILYQF EKRRRLYHLMKQSLNOPELLKQY
ITQLELE"
6482..7849
/locus_tag="SE0912"
6482..7849
/locus_tag="SE0912"
/codon_start=1
/transl_table=1
/product="fth protein"
/protein_id="AA04509.1"
/db_xref="GI:27315375"
/translation="NAFGLSDRLQATMOKRGKGVTEADITKMYREVLALLADV
NFRVYVEPVTYSERAIGSDMOSITGGQVIRKYOBELTSLMGKSTIKANKPPT
VYMGVLOGAGTTTACGLALIMRKYKKNKDLVADYIPRAIIDQLOTVGGQIDIPV
YSGDQVSPQOIVENALNKAKEHLPIIIDTAGRLHIDALNELQEKELSKDEI
MLVDMATGDDAVNVAQSFDRGLDSGTTLKLGQDGRGAAISRSVQKPTVGM
SEKLDLELHFERVASHNLLGNDVLSLIEKAAQDVQVQAKULEKKNSESSFTLDLF
LEQLDQVKNLGLDIDIMKIPGMKNMGGLKLNNSDQIDHIAIYIQTWPERSNPE
TLNVSRRKRIANGSGRSLQEVNRLKQFNDMKKMKQFTGGGKAKKRNQNMNLKG
MNLPF"
8281..8556
/locus_tag="SE0913"
8281..8556
/locus_tag="SE0913"
/codon_start=1
/transl_table=1
/product="30S ribosomal protein s16"
/protein_id="AA04510.1"
/db_xref="GI:27315376"
/translation="MAVIRITRGRSRNPFYRIVAADRSPRGRHIIEOQLQYPAH
VNAPEKIDEELATKWLHOGAKPTQVTHILSRGLIKKFDQKNK"
8683..9186
/locus_tag="SE0914"
8683..9186
/locus_tag="SE0914"
/codon_start=1
/transl_table=1

[illegible]

Db	296797	TATCATTAGCAATTCACAGATTATATCAAGTTTATTCACAGCTAAATACACAGCTTTCTTAG	296856		
Oy	181	ATTGTTGATAAATTTCAGGACCATTTTCATCGCTTTGTGTCAATTCCTTCACGACGTTAA	240		
Db	296857	ATTATTAAGCATTGTGTGCACCATTTCTGTCATTGTGTAACTCTTCTTAACTTGA	296918		
Oy	241	CCATTAATAATATATACATCCGCTTCATCATGCGCATTTTCGCTGCGCTTAATTTGTG	300		
Db	296917	CCATTAATAATGATGATCATGCTGCTTCATATGATTTCTGCTGCTGACGAGATTGAG	296978		
Oy	301	TTTAGATGGTGCATCACCAATTTCAATGACAGCTGTATCATATATATTTGAAATCATGT	360		
Db	296977	TTTGAAAGAGAGCATCTCCAAATTTCAATGACAGCTGTATCATATATGTTAAATTCATGAG	297038		
Oy	361	TTAACCATTGACCTGACGAATTAATAC	387		
Db	297037	TTAACCATCTGCCAGATGAATTAATAC	297063		
RESULT 8	AE017028/c	290117 bp	DNA	linear	BCT 30-APR-2003
LOCUS	Bacillus anthracis str. Ames	section 5 of 18	of the complete		
DEFINITION	genome.				
ACCESSION	AE017028	AE016879			
VERSION	AE017028.1	G1:30255149			
SOURCE	Bacillus anthracis str. Ames				
ORGANISM	Bacillus anthracis str. Ames				
REFERENCE	1 (bases 1 to 290117)				
AUTHORS	Read,T., Peterson,S., Tourasse,N., Baillie,L., Paulsen,I., Nelson,K., Tetteelin,H., Fouts,D., Eisen,J., Gill,S., Holtzapple,E., Ostred,O., Helgason,E., Riststone,J., Wu,M., Kolonay,J., Beaman,M., Dodson,R., Brlkac,L., Gilm,M., Deboy,R., Madupu,R., Daugherty,S., Durkin,A., Haft,D., Nelson,M., Peterson,J., Pop,M., Khouri,H., Radune,D., Benton,J., Mahmoud,Y., Jiang,L., Hance,I., Weidman,J., Berry,K., Plaut,R., Wolf,A., Watkins,K., Nierman,M., Hazen,A., Cline,R., Redmond,C., Thwaite,J., White,O., Salzberg,S., Thomson,B., Friedlander,A., Koehler,T., Hanna,P., Kolsto,A.-B. and Fraser,C.				
TITLE	The genome sequence of Bacillus anthracis Ames and comparison to closely related bacteria				
JOURNAL	Nature 423 (6935), 81-86 (2003)				
MEDLINE	22608414				
PUBMED	12721629				
REFERENCE	2 (bases 1 to 290117)				
AUTHORS	Read,T., Peterson,S., Tourasse,N., Baillie,L., Paulsen,I., Nelson,K., Tetteelin,H., Fouts,D., Eisen,J., Gill,S., Holtzapple,E., Ostred,O., Helgason,E., Riststone,J., Wu,M., Kolonay,J., Beaman,M., Dodson,R., Brlkac,L., Gilm,M., Deboy,R., Madupu,R., Daugherty,S., Durkin,A., Haft,D., Nelson,M., Peterson,J., Pop,M., Khouri,H., Radune,D., Benton,J., Mahmoud,Y., Jiang,L., Hance,I., Weidman,J., Berry,K., Plaut,R., Wolf,A., Watkins,K., Nierman,M., Hazen,A., Cline,R., Redmond,C., Thwaite,J., White,O., Salzberg,S., Thomson,B., Friedlander,A., Koehler,T., Hanna,P., Kolsto,A.-B. and Fraser,C.				
TITLE	Direct Submissio				
JOURNAL	Submitted (26-MAR-2003) The Institute for Genomic Research, 9712				
FEATURES	Medical Center Dr. Rockville, MD 20850, USA				
SOURCE	Location/Qualifiers				
	1. 290117				
	/organism="Bacillus anthracis str. Ames"				
	/mol_type="genomic DNA"				
	/db_xref="taxon:198094"				
	331..1929				
	/locus_tag="BA1197"				
	331..1929				
	/locus_tag="BA1197"				
	/note="similar to GP-107117143, and GP-107117143;"				

gene EAKFMGKCDIINTDREGIT*
7854.. .9680
/gene="pepf-1"

Query Match 54.4%; Score 210.6; DB 1; Length 290117;

Best Local Similarity 71.7%; Pred. No. 9,2e-38;
Matches 276; Conservative 0; Mismatches 109; Indels 0; Gaps 0;

OY 3 TCTCTCTCTCTTCAACCAATGAGAAACAAGTCACTAAGACCAAGCACTAA 62

DB 280879 TCTCTCTCTCTTGGAAAAAGCTGTCACCTTCATCTAATAGTCACTAATCTAAAT 280820

OY 63 CCATGTACCCGTGATATGCGATACGTTACCAAAATCTATGATGAATATACACG 122

DB 280819 CCGTGTACCTGAGATGGCAATGCGCGCAAAATCTATGATGAATATGAATG 280760

OY 123 TCTGTACGATTTCCATATATCTATCTTTGTAACCGCTATACGACCGGTTTTAGAT 182

DB 280759 TCACCTGCGATTTCCGATGTCACCTTATTTACTGCACTACAGCTGTTATTTGAA 280700

OY 183 TTGTATAAATTTGACGACATTTCAATCGCTTGTGTCAATCCCTCAGCAGCTAAC 242

DB 280699 CGATATATATTTTTCACACTCTTCATCCGCTGCTTACACCGTCACGACATTTGTC 280640

OY 243 ATAAAAATATAATCCGCTCATCTATGCGATTTCTGCGCTGCGCTAATTTGCTT 302

DB 280639 ATAAAAATATACATCTCTCATCAATTTGCTACTGCTTGTGACGAAATTTGTC 280580

OY 303 TGAATGTGATCACCACCAATTTCAATACCACTGTATCAATATATTTGAATCATGTGTT 362

DB 280579 AAAAAAGCTCGCTCCCAATATCAATTCACCTGTATCAATATATTTGAATCATGTGTT 280520

OY 363 AACCATTCACCTGAGAAATAAATAC 387

DB 280519 AACCATTCACCGCTATTAATAC 280495

RESULT 9
AE017002/c 304680 bp DNA linear BCT 16-MAY-2003

LOCUS AE017002 Bacillus cereus ATCC 14579 section 5 of 18 of the complete genome..

ACCESSION AE017002.1 GI:29894935

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

source

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

1..304680
/organism="Bacillus cereus ATCC 14579"

/mol_type="genomic DNA"

/strain="ATCC 14579"

/db_xref="ATCC:14579"

/db_xref="taxon:1226900"

68..427

/locus_tag="BC1239"

68..427

/locus_tag="BC1239"

/locus_tag="BC1239"

/locus_tag="BC1239"

/locus_tag="BC1239"

/locus_tag="BC1239"

/locus_tag="BC1239"

/locus_tag="BC1239"

/locus_tag="BC1239"

/locus_tag="BC1239"

/locus_tag="BC1239"

/locus_tag="BC1239"

/locus_tag="BC1239"

/locus_tag="BC1239"

/locus_tag="BC1239"

/locus_tag="BC1239"

/locus_tag="BC1239"

/locus_tag="BC1239"

/locus_tag="BC1239"

/locus_tag="BC1239"

/locus_tag="BC1239"

/locus_tag="BC1239"

/locus_tag="BC1239"

/locus_tag="BC1239"

/locus_tag="BC1239"

/locus_tag="BC1239"

/locus_tag="BC1239"

/locus_tag="BC1239"

/locus_tag="BC1239"

/locus_tag="BC1239"

1..304680
/organism="Bacillus cereus ATCC 14579"

/mol_type="genomic DNA"

/strain="ATCC 14579"

/db_xref="ATCC:14579"

/db_xref="taxon:1226900"

68..427

/locus_tag="BC1239"

68..427

/locus_tag="BC1239"

/locus_tag="BC1239"

/locus_tag="BC1239"

/locus_tag="BC1239"

/locus_tag="BC1239"

/locus_tag="BC1239"

/locus_tag="BC1239"

/locus_tag="BC1239"

/locus_tag="BC1239"

/locus_tag="BC1239"

/locus_tag="BC1239"

/locus_tag="BC1239"

/locus_tag="BC1239"


```

RBS      C5SDRYLDRSFPIWEINHLKLNKRNKREIDJLNGCY"
          complement(10356..10361)
gene      /gene="11n1792"
          10476..11787
RBS      /gene="gsaB"
          10476..10481
CDS       /gene="gsaB"
          10489..11787

```

/product="glutamate-1-semialdehyde aminotransferase"
 /protein_id="CAC97024.1"
 /db_xref="GI:16414296"
 /db_xref="SPRMBL:092A5"
 /translation="MDHSKSKLHDEALLHVGCVSPSKSNKCVGGGIPPTMEANAK
 AYFDYDKNKYIDYLAAPFPIITGHAHPHTTEATITKAQGVLYGPTTHLETFFAKMI
 KAIPILEKYEVRTNGTSEAVMTTIRVAAYATGRTIKFEGCNGHGFIDYLAACSGQV
 STGIDISAGVTRSAEIVTVPFNDLTSFEFALAVNGDVAALVEPIVGNNGNGLP
 AEGFLDAVLEHNGSLAIVTDEVITAFRFGGAAYGCLVIDLPAKMIIGGGLP
 GAVGGVDIMETVAPLCPAYOAGHAGPASTISGICAEVLQDEELVDRFVCSMI
 KEIDIEAALIKHICIAATVNOYCALTYVTDEDPRTNYREAGATNGDLGKFFGMGLEEBE
 INLAPSKTCAWMTTSAHSADLTETIAQVDFEKLQGR"
 join(11879..12981,12990..13016)
 /gene="lin1794"

Query Match	53.2%	Score 205.8	DB 1	Length 313450
Best Local Similarity	70.9%	Pred. No. 1.1e-36		
Matches 273; Conservative	0	Mismatches 112	Indels 0	Gaps 0

Db	276305	TCCTCTTCTCTTTTGGAAATGAGCAGCAACAGCATATACCATATCTACCAAGCCCTAGT	276366
Qy	63	CCATCTGACCCCGATATGCGGATACGGTTCACCAATCTCTAATGAATAGAAATCATACAGC	122
Db	276365	CCATGAGAACCGAAATTTGGATACGGCTCACCACCAACGAGAAATAAATCTCATTAATC	276422
Qy	123	TCGTAGCGCATTTCCATATTATCTACTTTGTTTAAACCGCTAAATACGACCGGTTTTAGAT	182
Db	276425	TGATCAGCGATTCTCGGTTTATCTACTTATTATTAATCGCTAAACAAATGGTTTATTAGAC	276488
Qy	183	TTGTATAAATTTTGAGCGACCACTTTCATCGCTTTGTGTCAATCCTTCACGCACTTAAC	242
Db	276485	CGGTAAAGATTTTGTCTACTGTCTGCTGCATGGTAACCCCTTCACGCACTTGGTA	276544
Qy	243	ATTAATAATATAACATCCGCTTCATCTATAGCGATTTCTGCTCGGCTCTAATTGTGT	302
Db	276545	ATTAATAATATACGCTCTGCTTCATCAATATGGGATTTGCGGTGGCGCGAATTTGGCT	276604
Qy	303	TGCAATGGTGCATCAGCAATTTTCATPACCACTGTATCAATATAATTGAATCATGTGT	362
Db	276605	MGAAATGGTTCGTCGGAAGAATCAATACCACTGTATCAATATGTTAAATCTTTTCCA	276664
Qy	363	AMCCATTCACTGGAAGATTAATAC	387
Db	276665	AGCCATTCCGGTGAATTAATATATGC	276689

AX413016/

LOCUS	AX413016	319630 bp	DNA	linear	PAT 02-SEP-2002
DEFINITION	Sequence 7 from Patent MO0228891.				
ACCESSION	AX413016				
VERSION	AX413016.1	GI:21445474			
KEYWORDS	.				
SOURCE	Listeria innocua				
ORGANISM	Listeria innocua				
REFERENCE	Bacteria; Firmicutes; Bacillales; Listeriaceae; Listeria.				
AUTHORS	Kunst, F. and Glaser, P.				
TITLE	Listeria innocua, genome and applications				
JOURNAL	Patent: WO 0228891-A 7 11-APR-2002;				
	INSTITUT PASTEUR (FR) ; CENTRE NATIONAL DE LA RECHERCHE				

BASE COUNT	ORIGIN
105207	a
55428	c
66726	g
92263	t
6	others

	Query Match	Similarity	53.2%;	Score	205.8;	DB	6;	Length	319630;	
	Best Local	Similarity	70.9%;	Pred.	No. 1.16-36;					
	Matches	273;	Conservative	0;	Mismatches	112;	Indels	0;	Gaps	0;
Qy	3	TCTTCTTCCTCTTCCACCAAAATGAGAAACCACTCATCTACACAGTCACCAAGACTTAA	62							
Db	51123	TCCTCTTCTCTTTTGGAAAAATGAGACACGAACACACTCAAGCATATACACCAAGCCCTAGT	5106							
Qy	63	CCATGTACCCCTATGTGCGATACGGTTTCACCAAAATCTTAATAAATCAATATACACG	122							
Db	51063	CCATGACAAACCAAAATGTGATACGGCTTCACCAAAACCAAGAAATTAAGTCATAAATTC	5100							
Qy	123	TCTGTAGGCAATTCACATATATATCTCTTGTAAACCCCTAAATACGACCGGTTTTTACAT	182							
Db	51003	TGATACAGCAATTCCTCGGTTATCTCTACTTATTAATCCCTAAACAATTTGGTTTATTAAC	5094							
Qy	183	TTGTATTAATAATTTGACGACCAATTTATCATCGCTTGTCTCAATCCTTCACGACCGTTAAC	242							
Db	50943	CGGTAAAGAAATTTTCTACTCTGTTGCTGCTCATCGGTAAACCCCTTCACGACCATTTGTA	5088							

	Db	50883	ATTAATAATTATTCAGCTGCTGTTCATCATCAATTTGGCAATTTCCGCTTGC GGGAATTTGCCCT	50822
	Oy	303	TGAATATGTCATACCAATTTCAATACCACCTGTATCAATTAATTTGAAATGTC TT	362
	Db	50823	AAGAATGTCCTCGCGGAAAAGATCAATACCACCTGTATCAATTAATTTGTAATTTCTTTTCCA	50766
	Oy	363	AACCATTCACCTGAGAGATAAATATAC	387
	Db	50763	AGCCATTCGCGTGATATATATATG C	50739
RESULT 12				
AL591981				
LOCUS		347050 bp	DNA	linear, BCT 06-JUN-2002
DEFINITION		Listeria monocytogenes strain EGD,	complete genome, segment 9/12.	
ACCESSION		AL591981		
VERSION		AL591981.1	GI:16411141	
KEYWORDS				
SOURCE		Listeria monocytogenes		
ORGANISM		Listeria monocytogenes		
REFERENCE		Bacteria; Firmicutes; Bacillales; Listeriaceae; Listeria.		
AUTHORS		Glaser, P., Frangeul, L., Buchrieser, C., Rusnok, C., Amend, A., Bacquer, P., Berche, P., Blocher, H., Brandt, P., Chakraborty, T., Charbit, A., Chetouiari, F., Couve, E., de Lencastre, A., Deboux, P., Domane, E., Dominguez-Bernal, G., Duchand, E., Durant, L., Dussaugel, O., Enliam, K.D., Esnlin, H., Portillo, F.G., Garrido, P., Gaullier, L., Goebel, W., Gomez-Lopez, N., Halin, T., Hauf, J., Jackson, D., Jones, L.M., Kaerst, U., Kretz, J., Kuhn, M., Kunst, F.,		

termInator
gene
terminator
CDS

INCUNFTAPHPVOEKKEFAERLCKKMKPKYTPVPKIIKIFILGEMAMTILDSORAY
EXLMSNHFEFRFETLOEALDLDL
5681..5708
/gene="lmo1694"
complement(join(5682..8291,8299..8304))
/gene="lmo1695"
complement(5682..5709)
/gene="lmo1695"
complement(5694..8291)
/gene="lmo1695"
/note="similar to putative membrane proteins"
/codon_start=1
/transl_table=11
/protein_id="CAC99773.1"
/db_xref="GI:1641149"
/db_xref="SPTREMBL:O8Y619"
/translacion="MKELMQAYAMFQKNSVAVIYFEFVNAVIFELINATCIDY
PSLKENTLOSPEOFIMFIVGLAVTPMLLYOVIVKLPKFSFVSHTVAGMITNT
FTNIGFGVGLGASRASFGKNSHKEITLAIKALFVGLSGLSVLSTLILIP
GFADHPVYHMLAGLCLYPLFTITKMSKSLFYDLPITKRELIITLILKMPAF
GCEPILITLKEBPDIKFPVPLFYIASVIGIASKVPQCTPVPVITLISQLSVSE
LALAMLTIRITFTYIIPVGLTFVQKAGKNDLEGLPLFLQKVAHRELVITY
GSGLLILISSAPNATIHVPFLKIMPENLEFSQITVAFGLLGLANGLECKTK
AYIIVIVGCAIFNTLAVFSMKAIFGLIQLFELANEFYREKLVYTSKVID
SIFIVCLAGYIVIGIYNSPNIKSKSEIPYLRIASEHMLVGVFVAVSLVYIY
IYLSTTKEKLSPEAVVAREHLAKMGNEVSHMFLROKLFMAEGELFSYRILA
DKMVLNGEPTGMOKMEALIEVQMNADRGYRPFYEVAGTMIPLYLHDGDFETIA
EDGFVQNFTHSGKKKGERALANKLERGTYFEIIEPENNIDETTLTILAVSDEMID
GREKFGSLGDFPDYVLEOAPITAKNGESTIYFASKVPSTDEMTSIDLRYSSEA
PSCIMDPLPIPLPEKAKEDGQTFNAGMPLAVNGSEKVFLEBERGLVYRSCFY
GPKGRNRSKYVTYEMDKFVAFPKRSIATFHLQMLILVKKRPLANSQVLDPELE
ETKKPDESE

RBS
gene
CDS

complement(8299..8304)
/gene="lmo1695"
complement(8318..9007)
/gene="lmo1696"
complement(8318..8992)
/gene="lmo1696"

Query Match 53.2% Score 205.8; DB 1; Length 347050;
Best Local Similarity 70.9%; Pred. No. 1,1e-36;
Matches 273; Conservative 0; Mismatches 112; Indels 0; Gaps 0;

QY 3 TCTCTCTCTTACCAAAATGAGAAACATCTCATCTACACCAAGACCTAA 62
DB 258933 TCCTCTCTCTTGGAAAATGAGCGGCAACACATCAAGCATGCGCAACCCCTAGT 258992
QY 63 CCATGAGCCCTGATATCGATACGGTTCACCAATCCTATGAAATGAAATCATACAG 122
DB 258993 CCATGAGCAAGAAATGGACGGCTCACCAACCAAGAAATGAAATGAAATTT 259052
QY 123 TCTGTACGATTCATATCTACTTGTAAACCGCTAATACGACCGTTTTTAAAT 182
DB 259053 TGTATCGCATTTCTGGGTTATCTATTAATCCCTAAACAAATGGTTATTAAGAC 259112
QY 183 TTGTATTAATTTGAGCGACATTTTCATCGCTTGTGTCAATCCCTTCCACGACGTTAAC 242
DB 259113 CGGTAAAGAAATTTTGTACTGTCTGCTCATCGGTAACCTCTTACAGACATTGGTA 259172
QY 243 ATTAATAATTAATACCGCTCATCTATGGGCAATTCGCTGCGCTGCTAAATTTGGTT 302
DB 259173 ATTAATAATTAATACGCTTCTGCTTCAATTCGCAATTCGCTTGTGCGCAATTTCTCT 259232
QY 303 TGGATGTGTCATACCAATTTCAATACCACTGTATCAATATATTAATGAAATCATGTGT 362
DB 259233 AAAAATGTTTATCGAAAGATTAACACCTGTATCAATATATGTTAAATTTCTTTTCCA 259292
QY 363 AACCATTCACCTGAAGAAATTAATAC 387
DB 259293 AGCCATTCGCTGAATTAATATATGC 259317

RESULT 13

AX417046
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
SOURCE

AX417046 349980 bp DNA linear PAT 02-SEP-2002
Sequence 4037 from Patent WO0228891.
AX417046
AX417046.1 GI:21449656
Listeria innocua
Listeria innocua
Bacteria; Firmicutes; Bacillales; Listeriaceae; Listeria.
Kunst, F. and Glaser, P.
Listeria innocua genome and applications
Patent: WO 0228891-A 4037 11-APR-2002;
INSTITUT PASTEUR (FR) ; CENTRE NATIONAL DE LA RECHERCHE
SCIENTIFIQUE (CNRS) (FR)
Location/Qualifiers
1. 349980
/organism="Listeria innocua"
/mol_type="genomic DNA"
/db_xref="taxon:1642"
/note="seq 2058, original length: 3,011,208 replaced
by-seq 2058: 0,000,001 to 0,349,980-seq 4032: 0,300,001 to
0,649,980-seq 4033: 0,600,001 to 0,949,980-seq 4034:
0,900,001 to 1,249,980-seq 4035: 1,200,001 to
1,549,980-seq 4036: 1,500,001 to 1,849,980-seq 4037:
1,800,001 to 2,149,980-seq 4038: 2,100,001 to
2,449,980-seq 4039: 2,400,001 to 2,749,980-seq 4040:
2,700,001 to 3,049,980-seq 4041: 3,000,001 to 3,011,208"
101055 a 72969 c 60688 g 115268 c
ORIGIN

Query Match 53.2% Score 205.8; DB 6; Length 349980;
Best Local Similarity 70.9%; Pred. No. 1,1e-36;
Matches 273; Conservative 0; Mismatches 112; Indels 0; Gaps 0;

QY 3 TCTCTCTCTTACCAAAATGAGAAACATCTCATCTACACCAAGACCTAA 62
DB 26305 TCCTCTCTCTTGGAAAATGAGCGGCAACACATCAAGCATGCGCAACCCCTACT 26364
QY 63 CCATGAGCCCTGATATCGATACGGTTCACCAATCCTATGAAATGAAATCATACAG 122
DB 26365 CCATGAGCAAGAAATGGACGGCTCACCAACCAAGAAATGAAATGAAATTT 26424
QY 123 TCTGTACGATTCATATCTACTTGTAAACCGCTAATACGACCGTTTTTAAAT 182
DB 26425 TGTATCGCATTTCTGGGTTATCTATTAATCCCTAAACAAATGGTTATTAAGAC 26484
QY 183 TTGTATTAATTTGAGCGACATTTTCATCGCTTGTGTCAATCCCTTCCACGACGTTAAC 242
DB 26485 CGGTAAAGAAATTTTGTACTGTCTGCTCATCGGTAACCTCTTACAGACATTGGTA 26544
QY 243 ATTAATAATTAATACCGCTCATCTATGGGCAATTCGCTGCGCTGCTAAATTTGGTT 302
DB 26545 ATTAATAATTAATACGCTTCTGCTTCAATTCGCAATTCGCTTGTGCGCAATTTCTCT 26604
QY 303 TGGATGTGTCATACCAATTTCAATACCACTGTATCAATATATTAATGAAATCATGTGT 362
DB 26605 AAGAAATGTCGCGGAAAGATTAACACCTGTATCAATATATGTTAAATTTCTTTTCCA 26664
QY 363 AACCATTCACCTGAAGAAATTAATAC 387
DB 26665 AGCCATTCGCTGAATTAATATATGC 266689

RESULT 14
AX641670
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

AX641670 349980 bp DNA linear PAT 21-FEB-2003
Sequence 2860 from Patent WO0101118.
AX641670.1 GI:28474431
Listeria monocytogenes
Listeria monocytogenes

/codon_start-1
/transl_table-11
/protein_id="AAC83944.1"
/db_xref="GI:1146197"
/translation="MKVKKLVVSMLSIAFVLLNLNFPFGLPDYLKIDFSVPAII
ALILYPLAGIAVEAIKKNVLOYIIOGSMAGVPGVAMFAGLTFTLPTPAFLFKLNS
AKGLAVSLILGTAANTILMSILNLTLLTPAYTPHSPALSDALKTAVAVACILPDM
IKGIVTVVFSILFIKLKPMIEOORSAH"
2686..2721
/note="putative"
/citation-[3]
complement(2726..2974)
/gene="fer"
complement(2726..2974)
/gene="fer"
/note="ypba: similar to B.stearotherophilus ferredoxin;
putative"
/citation-[3]
/codon_start-1
/transl_table-11
/product="ferredoxin"
/protein_id="AAC83945.1"
/db_xref="GI:1146198"
/translation="MAKTVIDKDTIACGACGAAADIVYDDEGIAFVTLDEKGV
VEVPEVLEEDMIDAFEGCPTDSIKVADEPFGDPLKE"
3240..4298
/gene="ypbB"
3240..4298
/gene="ypbB"
/note="putative"
/citation-[3]
/codon_start-1
/transl_table-11
/protein_id="AAC83946.1"
/db_xref="GI:1146199"
/translation="MSLNYFDIMVLDLCKINGERSGSAVFLFKGRSSQTOTDAGL
FOTAKYFGMAKCSRSDISADLLEKHSYLAPMSSEDYKYTAAGAAVLRALSERP
MPVHGAGHYOAGVLAHKLRLSLVYLSHKQSGSHQYIPTKDKHTLHWKYSRH
TDHIEAKSLFAMLELAHLKRIENKAQIPYSLTSHRIQTSROLADTLKDEENVY
IMFNASVHYFIHSLPCECDALIKDLSVHLDMALTESTRTKTMQWKGQEPRIORAEI
RLKATATIEDHIVEISLHEPAPMINDVVSADQLIAEFAKRMRTKIKINGIDGLEOR
FSYQIRALAKVOQYD"
4291..5781
/gene="recs"
4291..5781
/gene="recs"
/function="DNA repair and homologous recombination"
/note="similar to Escherichia coli recQ gene product;
putative"
/citation-[3]
/codon_start-1
/transl_table-11
/product="DNA or RNA helicase, DNA-dependent ATPase"
/protein_id="AAC83947.1"
/db_xref="GI:1146200"
/translation="MTRLOQTLYOFQFTSPFKGQODIIESILSGKDTTAMLPITGGGK
SLCYOLPGYMLDGNVLIVSLSLMEQVOQLKAREKRAAALNSMLNROEQVLEH
IHRKFTLSPALQSPYVLEKLSYVLSFVIDEACHISEMGRDPRPSLQAKRI
KLHPPVLAATATATKETLQDVNHLLEQHAVRHLNSVRPNIALVEENAATGAKID
RYIOLVENLOGPGIVCPTPRMAKELAGEIKSTSSRADPYGSGESGRLLIIOOFI
HNOLDVICCTNAFGMGVDPDIRYVIFHLPORTAEPMOEIGRAGDCKPSYILSRA
PGDELOBOITOMESYAEELADIVIRVLTETTERDPRRLPVLLQYGCVEYERARHII
LPNGKTSVELAKREISYRBELEKKNRHSFLQDCCRLQROLTLTYPESTEPDGN
LPCSHCGFDLSLEQNGERSKMAPIDSNSELRHIFSLQTVGEIN"
5841..6410
/gene="ypbD"
5841..6410
/gene="ypbD"
/note="putative"
/citation-[3]
/codon_start-1
/transl_table-11
/protein_id="AAC83948.1"

/db_xref="GI:1146201"
/translation="MTDLITFPAAGACLFEEEDRVYLKNDIRMTIIMYGVIAV
IGVLADMAVMKFPSPHLYDDGINKRIKFSKSIHIIIFLULAFAEEDLFGVQLQTH
IYLTASLIFPALHFRYLSKWLFIWVAISFLLGIMYEMTGNLFPVMAHPITDAVF
ACQIRFEHVRROKHDEHVESREKSPESL"
6361..7083
/gene="ypbE"
6361..7083
/gene="ypbE"
/note="putative"
/citation-[3]
/codon_start-1
/transl_table-11
/protein_id="AAC83949.1"
/db_xref="GI:1146202"
/translation="MTNMSRVBRKRAONLYEDQNAALADYVDGSESLPTROSVKNOR
EOKKQKTKTPELTVLAVIFVFPVIVLTVLYLKSHPDNHDDYEDVEIDSSQSKYE
VYKSEEDKHDTADTKETALOKESKKEPEDSKPKEOTAAKDKOTAVAKESDKNKEAT
AAASSTOSTVOOQOPAEVONPNRVKHTVOKKETLVIRISKYKNSRTGEKIRA
YHNLKNDYVTCQVLDIPLMDE"
7146..7589
/gene="ypbF"
7146..7589
/gene="ypbF"
/note="putative"
/citation-[3]
/codon_start-1
/transl_table-11
/protein_id="AAC83950.1"
/db_xref="GI:1146203"
/translation="MESTANOLDQFTDAPTROMLQVLRKQFENYAAOCRRMKAS
LITGLGLQVMINKSPFLTDLDELSHTFTLPMULTATATYSTFKKREKSETD
FHLRCEIIVKSTDLMPQDKMKARESVFHMKKHYDINLFESK"
7597..7619
/note="putative"
/citation-[3]
/codon_start-1
/transl_table-11
/protein_id="AAC83951.1"
/db_xref="GI:1146204"
/translation="MKSIVKIAGVLTAAAMATPAKNVATPAKGNHLLKHTHPPLSKKRG
PLTFEISDILHRLIDODLEKASHAPHLVIGGLAGGVSPSARIEENIKRLVHF
GVPIVFMGNNDSEYVROHKLYSIFRAGVITLIRNESVPSYNGHTIANGVDIDREM
DHIEBKIRLDESQNLIVCHNPHEIHQINEDDGIDVYIISGHTHGOIRFGFQYEL
51.3%: Score 198.4; DB 1: Length 24887;
Best Local Similarity 71.2%: Pred. No. 6.7e-35;
Matches 262; Conservative 0; Mismatches 106; Indels 0; Gaps 0;
Query Match
20 AAAATGAGAAACATGCATCTAACACGACCAAGACCTAACCATGTGACCTGATAT 79
21672 AAAATGCTCTGCAAGCGATCCAGTAATATCACCCAGTCGCGTGTCCGCAAT 21613
80 CGGATACGGTTCACCAATTCCTAATGATTAATATCAACGCTGTCACGATTTCCAT 139
21612 TGGATACGGCTGCGCAACACCTACGAAATTAATTAATATGCTGCTCTCTTCGT 21553
140 ATTATCTACTTGTTAACCGCTAATATACGACCGCTTTTATGATTTGTATTAATTTGAC 199
21552 GTTATCAGTATTATTAACCGCTTAATTAATTAATTAATTAATTTTTCGC 21493
200 GACCATTCATCGCTTGTGTCATCTTCAGCAGCTTAACCATTAATTAATTAATGATC 259
21492 CACTTTCATCAGCAGCTGTCAGCGCTTCACGGCGCTTACCATTAATTAATTAATGATC 21433
260 CGCTTCATCTATGCGGATTTTGCCTGCGCTTAATTTGTTTGAATGCTGATCATC 319

Sat Sep 13 12:47:59 2003

us-09-815-242-1463.rge

Page 19

Db 21432 CGCTTCATCCATGGCGATTTCAGCTTGCCTGGCGCAATCTGCGCTAAAAACGGCTCATCACC 21373
Oy 320 AATTTCATACACACCTGTATCATTAATTAATTGAATCATGTGTTAACCATTCACCTGAGAGA 379
Db 21372 GATATCAATACCGCGCGGTATCATCAAAATTAATCAATATTCAGGCATTCAGCCGAGCT 21313
Oy 380 ATAAATAC 387
Db 21312 GTATATCC 21305

Search completed: September 12, 2003, 19:27:28
Job time : 1589.02 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 12, 2003, 17:16:39 ; Search time 438.085 Seconds
(without alignments)
8041.284 Million cell updates/sec

Title: US-09-815-242-4228

Perfect score: 1305

Sequence: 1 atgactaaccttagtagc.....atattatagctcgaagaga 1305

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N.Geneseq.19Jun03.*

1: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT.*
2: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT.*
3: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT.*
4: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT.*
5: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT.*
6: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT.*
7: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT.*
8: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT.*
9: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT.*
10: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT.*
11: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT.*
12: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT.*
13: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT.*
14: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT.*
15: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT.*
16: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT.*
17: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT.*
18: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT.*
19: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.*
20: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.*
21: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.*
22: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.*
23: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.*
24: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT.*
25: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1305	100.0	1305	23 AAS51646	Staphylococcus aur
2	1305	100.0	1311	23 AAS54865	Staphylococcus aur
3	1292.2	99.0	1311	22 AAF8461	Staphylococcus aur
4	1290.6	98.9	1311	23 AAS54997	Staphylococcus aur
5	1230.2	94.3	3621	18 AAV74669	Staphylococcus aur
6	921	70.5	1312	24 ABA00883	Staphylococcus epi
7	919.4	70.5	3269	22 AAF54708	S. epidermidis gen
8	683.2	52.4	2944528	24 ABA03041	Listeria monocytog

9 676.8 51.9 319630 24 ABO67194
10 676.8 51.9 3011208 24 ABO69245
11 667.6 51.2 960 22 AAS53683
12 600 46.0 1308 24 ABA68457
13 600 46.0 2155561 24 ABA71527
14 599.4 45.9 1311 24 ABA75008
15 582.4 44.6 2365589 24 ABA90521
16 550.4 42.2 1308 25 ABA07474
17 550.4 42.2 1311 21 AAS54516
18 550.4 42.2 1311 21 AAS54516
19 550.4 42.2 3066 19 AAV52212
20 550.4 42.2 2162598 25 AAS56454
21 549.6 42.1 1308 24 ABA68458
22 383.8 29.4 387 23 AAS48886
23 367.2 28.1 372 23 AAS50706
24 336.8 25.8 2791 24 ABO70805
25 318.4 24.4 738 23 AAS55783
26 308.6 23.6 675 23 AAS51384
27 308.6 23.6 678 23 AAS52955
28 296.4 22.7 298 23 AAS50205
29 296.4 22.7 298 23 AAS50723
30 293.6 22.5 580073 18 AAT58840
31 270.8 20.8 33140 22 AAF28536
32 241.8 18.5 640681 24 ABA92787
33 241.2 18.5 246 23 AAS49268
34 210.4 16.1 1512 22 AAF94379
35 210.4 16.1 1512 24 ABA64943
36 210.4 16.1 1515 23 AAS53235
37 208.8 16.0 1830121 17 AAT42063
38 203.4 15.6 78845 21 AAB81463
39 203.4 15.6 349980 21 AAF21608
40 203.4 15.6 1437668 21 AAB81490
41 198.6 15.2 910715 20 AAX20248
42 190.6 14.6 1500 25 ABA239377
43 190 14.6 1512 23 AAS52517
44 190 14.6 11574 22 AAS46244
45 178.6 13.7 1557 22 AAB66542

ALIGNMENTS

RESULT 1

AAS51646
ID AAS51646 standard; DNA: 1305 BP.

XX AC AAS51646;

XX DT 13-FEB-2002 (first entry)

XX DE Staphylococcus aureus DNA for cellular proliferation protein #63.

XX KW Antisense; ds: prokaryotic cellular proliferation gene;

XX KW antibiotic; antibacterial; drug design.

XX OS Staphylococcus aureus.

XX PN WO200170955-A2.

XX PD 27-SEP-2001.

XX PF 21-MAR-2001; 2001WO-US09180.

XX PR 21-MAR-2000; 2000US-191078P.

XX PR 23-MAY-2000; 2000US-206848P.

XX PR 26-MAY-2000; 2000US-207727P.

XX PR 23-OCT-2000; 2000US-242578P.

XX PR 27-NOV-2000; 2000US-253625P.

XX PR 22-DEC-2000; 2000US-257931P.

XX PR 16-FEB-2001; 2001US-269308P.

XX XX (ELIT-) ELITRA PHARM INC.

XX XX

PI Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;
 PI Yamamoto RT, Xu HH;
 XX WPI: 2001-611495/70.
 DR P-PSDB: AAU33787.
 XX New polynucleotides for the identification and development of
 PT antibiotics, comprise sequences of antisense nucleic acids -
 XX
 PS Claim 27: Seq ID No 4228; 51pp; English.
 XX The invention relates to antisense inhibitors of genes essential to
 CC prokaryotic cellular proliferation, their use in identifying the
 CC genes, their use in the discovery of novel antibiotics, the essential
 CC genes themselves and the encoded proteins. The prokaryotes used are
 CC *Escherichia coli*, *Staphylococcus aureus*, *Salmonella typhi*, *Klebsiella*
 CC *pneumoniae*, *Pseudomonas aeruginosa* and *Enterococcus faecalis*. The
 CC invention is also useful for the identification of potential new targets
 CC for antibiotic development. The antisense nucleic acids can also be used
 CC to identify proteins used in proliferation, to express these proteins,
 CC and to obtain antibodies capable of binding, to the expressed proteins.
 CC The proteins can be used to screen compounds in rational drug discovery
 CC programmes. The antisense nucleic acid sequence is also useful to screen
 CC for homologous nucleic acids which are required for cell proliferation in
 CC a wide variety of organisms. The present sequence encodes an
 CC essential prokaryotic cellular proliferation protein.
 CC Note: The sequence data for this patent did not form part
 CC of the printed specification, but was obtained in electronic
 CC format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences.
 XX
 SQ Sequence 1305 BP; 447 A; 181 C; 277 G; 400 T; 0 other;

Query Match 100.0%; Score 1305; DB 23; Length 1305;
 Best Local Similarity 100.0%; Pred. No. 5.2e-290;
 Matches 1305; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ATGCTAAACCTATAGTACCTATTTAGTGTAGGTAGGCTTAATGTAGGTAAATCTACATTTT 60
 DB |||||||
 OY 1 ATGCTAAACCTATAGTACCTATTTAGTGTAGGTAGGCTTAATGTAGGTAAATCTACATTTT 60
 DB |||||||
 OY 61 AATAGAAATAGTTGGAGAACGTTGTTCCGATTTGTTGAAGATACGCCAGGTGTAACACGAGAT 120
 DB |||||||
 OY 61 AATAGAAATAGTTGGAGAACGTTGTTCCGATTTGTTGAAGATACGCCAGGTGTAACACGAGAT 120
 DB |||||||
 OY 121 CGTATTTATCTTCAGGTGAGTGTGTTAACACATGATTTCAATATTTATGATACAGGTGTT 180
 DB |||||||
 OY 121 CGTATTTATCTTCAGGTGAGTGTGTTAACACATGATTTCAATATTTATGATACAGGTGTT 180
 DB |||||||
 OY 181 ATGAAATAGTTGATGACCATTCACCAATTCACCAATTCACCAATTCACCAATTCACCAATTC 240
 DB |||||||
 OY 181 ATGAAATAGTTGATGACCATTCACCAATTCACCAATTCACCAATTCACCAATTCACCAATTC 240
 DB |||||||
 OY 241 GATGAAGCGGATGTTATTTATTTATGTTTAACTGCGTGAAGGATTGACACAAAGCGAT 300
 DB |||||||
 OY 241 GATGAAGCGGATGTTATTTATTTATGTTTAACTGCGTGAAGGATTGACACAAAGCGAT 300
 DB |||||||
 OY 301 GAATGGTTCGCTCAAAATTTATACAAATCTAAAAACCGGTGCTATTAGCGGTTAACAAA 360
 DB |||||||
 OY 301 GAATGGTTCGCTCAAAATTTATACAAATCTAAAAACCGGTGCTATTAGCGGTTAACAAA 360
 DB |||||||
 OY 361 GTAGATAATATGGAATTCGCTACAGACGCTGATGATTTCTATTATTAGGATTTGGTGAA 420
 DB |||||||
 OY 361 GTAGATAATATGGAATTCGCTACAGACGCTGATGATTTCTATTATTAGGATTTGGTGAA 420
 DB |||||||
 OY 421 CCGTATCAATATCAGGTTCACATGTTTATAGGCTTGGTCACTGTTAGATCCATGTT 480
 DB |||||||
 OY 421 CCGTATCAATATCAGGTTCACATGTTTATAGGCTTGGTCACTGTTAGATCCATGTT 480
 DB |||||||
 OY 481 TCTCATTTTGGTGAAGAGGAAGATTCCTTTATGATGAAGATACAAATTCGACTATCCATT 540
 DB |||||||
 OY 481 TCTCATTTTGGTGAAGAGGAAGATTCCTTTATGATGAAGATACAAATTCGACTATCCATT 540
 DB |||||||

OY 541 ATTTGGACGACCAACGCTAGGTAAATCAAGTTTAACTGCTATTTAGTGTGAAGATCCG 600
 DB |||||||
 OY 541 ATTTGGACGACCAACGCTAGGTAAATCAAGTTTAACTGCTATTTAGTGTGAAGATCCG 600
 DB |||||||
 OY 601 GTTATCGTTTCTAAATGTTTGCAGGACAAACGAGAGACGCTATTGATACAGAGTATATTTAT 660
 DB |||||||
 OY 601 GTTATCGTTTCTAAATGTTTGCAGGACAAACGAGAGACGCTATTGATACAGAGTATATTTAT 660
 DB |||||||
 OY 661 GATGGACAAAGATTTATGTTTAACTGCTGCTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 720
 DB |||||||
 OY 661 GATGGACAAAGATTTATGTTTAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 720
 DB |||||||
 OY 721 GAATCAACTGACAAATATTCAGTATTAAAGCGCTTTAAAGCGGATTTGAAGCTTCCAAATGTT 780
 DB |||||||
 OY 721 GAATCAACTGACAAATATTCAGTATTAAAGCGCTTTAAAGCGGATTTGAAGCTTCCAAATGTT 780
 DB |||||||
 OY 781 GTTTTATGCTGTTATTTGATGCGAACAAGGCAATTTATGCAACAAGATAAAGCTGTTGCGAGA 840
 DB |||||||
 OY 781 GTTTTATGCTGTTATTTGATGCGAACAAGGCAATTTATGCAACAAGATAAAGCTGTTGCGAGA 840
 DB |||||||
 OY 841 TATGCAATGAAACAGGTAAAGCAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 900
 DB |||||||
 OY 841 TATGCAATGAAACAGGTAAAGCAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 900
 DB |||||||
 OY 901 AAGATAGTAAACGATGAAGAAATTTGCAAGATCAAGTACGTAAGCAATTTCCAAATTTTAA 960
 DB |||||||
 OY 901 AAGATAGTAAACGATGAAGAAATTTGCAAGATCAAGTACGTAAGCAATTTCCAAATTTTAA 960
 DB |||||||
 OY 961 GATTATGCAACAATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1020
 DB |||||||
 OY 961 GATTATGCAACAATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1020
 DB |||||||
 OY 1021 CCTTACATTAATGAAGCAAGTGAACCAATTAAGCAAGTGAACCAATTAAGCAAGTGAACCAATTA 1080
 DB |||||||
 OY 1021 CCTTACATTAATGAAGCAAGTGAACCAATTAAGCAAGTGAACCAATTAAGCAAGTGAACCAATTA 1080
 DB |||||||
 OY 1081 GAAGTCTTACTGATGCAATTTCCATGACCCCTACACCAACGACACAAAGGTAGACGTTG 1140
 DB |||||||
 OY 1081 GAAGTCTTACTGATGCAATTTCCATGACCCCTACACCAACGACACAAAGGTAGACGTTG 1140
 DB |||||||
 OY 1141 AATGCTTTTATGCAACAAAGTTGCTGCTATAGAACCCGACATTTGCTGCTGCTGCTGCTGCTGCT 1200
 DB |||||||
 OY 1141 AATGCTTTTATGCAACAAAGTTGCTGCTATAGAACCCGACATTTGCTGCTGCTGCTGCTGCTGCT 1200
 DB |||||||
 OY 1201 GATGTAGATTAATGCAATTTTCTTATAAAGGCTATTAGAGATCAAAATCCGTCGCGCT 1260
 DB |||||||
 OY 1201 GATGTAGATTAATGCAATTTTCTTATAAAGGCTATTAGAGATCAAAATCCGTCGCGCT 1260
 DB |||||||
 OY 1261 TTTGGTGTGAAAGGTACACCAATTCATATTATAGCTCGAAAGAGA 1305
 DB |||||||
 OY 1261 TTTGGTGTGAAAGGTACACCAATTCATATTATAGCTCGAAAGAGA 1305
 DB |||||||

RESULT 2
 AAS4865
 ID AAS4865 standard; DNA; 1311 BP.
 XX
 AC AAS4865;
 XX
 XX
 DT 13-FEB-2002 (first entry)
 XX
 XX
 DE Staphylococcus aureus DNA for cellular proliferation protein #1177.
 DE
 KW Antisense; ds; prokaryotic cellular proliferation gene;
 KW antibiotic; antibacterial; drug design.
 XX
 OS Staphylococcus aureus.
 XX
 PN WO200170955-A2.
 XX
 PD 27-SEP-2001.
 XX
 XX 21-MAR-2001; 2001WO-US09180.

XX 21-MAR-2000; 2000US-191078P.
 PR 23-MAY-2000; 2000US-206848P.
 PR 26-MAY-2000; 2000US-207272P.
 PR 23-OCT-2000; 2000US-242578P.
 PR 27-NOV-2000; 2000US-253645P.
 PR 22-DEC-2000; 2000US-257931P.
 PR 16-FEB-2001; 2001US-269308P.
 XX (ELIT-) ELITRA PHARM INC.
 XX
 XX Haselbeck R, Ohlson KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;
 PI Yamamoto RT, Xu HH;
 XX WPI: 2001-611495/70.
 DR P-P5DB: AAU37006.
 PT
 PT New polynucleotides for the identification and development of
 PT antibiotics, comprise sequences of antisense nucleic acids -
 XX
 XX Claim 27: Seq ID No 8502; 511pp; English.
 XX
 CC The invention relates to antisense inhibitors of genes essential to
 CC prokaryotic cellular proliferation, their use in identifying the
 CC genes, their use in the discovery of novel antibiotics, the essential
 CC genes themselves and the encoded proteins. The prokaryotes used are
 CC Escherichia coli, Staphylococcus aureus, Salmonella typhi, Klebsiella
 CC pneumoniae, Pseudomonas aeruginosa and Enterococcus faecalis. The
 CC invention is also useful for the identification of potential new targets
 CC for antibiotic development. The antisense nucleic acids can also be used
 CC to identify proteins used in proliferation, to express these proteins,
 CC and to obtain antibodies capable of binding to the expressed proteins.
 CC The proteins can be used to screen compounds in rational drug discovery
 CC programmes. The antisense nucleic acid sequence is also useful to screen
 CC for homologous nucleic acids which are required for cell proliferation in
 CC a wide variety of organisms. The present sequence encodes an
 CC essential prokaryotic cellular proliferation protein.
 CC Note: The sequence data for this patent did not form part
 CC of the printed specification, but was obtained in electronic
 CC format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences.
 XX
 SQ Sequence 1311 BP; 451 A; 181 C; 277 G; 402 T; 0 other;
 Query Match 100.0%; Score 1305; DB 23; Length 1311;
 Best Local Similarity 100.0%; Pred. No. 5.2e-290;
 Matches 1305; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 ATGACTAACTATAGTAGCTATTGTAGGTAGGCGCTTAATGTAGGTAATCTACAAATTTT 60
 DB 1 ATGACTAACTATAGTAGCTATTGTAGGTAGGCGCTTAATGTAGGTAATCTACAAATTTT 60
 QY 61 AATGAGTAGTGGAGAACGTGTTTCGATTGTTGAAGATACCCAGGTGTACACGAGAT 120
 DB 61 AATGAGTAGTGGAGAACGTGTTTCGATTGTTGAAGATACCCAGGTGTACACGAGAT 120
 QY 121 CGTATTATTTCTTCAGGTAGTGGTTAAACACATGATTTCATATTATTGATACAGGTGGT 180
 DB 121 CGTATTATTTCTTCAGGTAGTGGTTAAACACATGATTTCATATTATTGATACAGGTGGT 180
 QY 181 ATTCAAATTCGTGATCCATTCACCAACACAAATAGAGCGCAGCGAATGCCNATA 240
 DB 181 ATTCAAATTCGTGATCCATTCACCAACACAAATAGAGCGCAGCGAATGCCNATA 240
 QY 241 GATGAACGGATGTTATTATTTTATGTTAACTGCGTGAAGGATTGACACAAAGCGAT 300
 DB 241 GATGAACGGATGTTATTATTTTATGTTAACTGCGTGAAGGATTGACACAAAGCGAT 300
 QY 301 GAAATGGTGGCTCAAAATTTTATACAAATCTAAAAACCGGTGCTATTAGCGGTTAACAAA 360
 DB 301 GAAATGGTGGCTCAAAATTTTATACAAATCTAAAAACCGGTGCTATTAGCGGTTAACAAA 360
 QY 361 GTAGATAATATGGAATGCGGTACAGACGTGTATGATTCTATTATAGGATTTGGTGAA 420

DB 361 GTAGATAATATGGAATGCGGTACAGACGTGTATGATTCTTATTCATTAGGATTTGGTGAA 420
 QY 421 CCGTATCCAATATCAGGGTCAATGTTAGTCTTGGTGACTTGTAGATCAGTGTGTT 480
 DB 421 CCGTATCCAATATCAGGGTCAATGTTAGTCTTGGTGACTTGTAGATCAGTGTGTT 480
 QY 481 TCTCATTTTGGTGAAGAGAGAGATCCTTATGATGAAGATACAAATTCGACTATCCATT 540
 DB 481 TCTCATTTTGGTGAAGAGAGAGATCCTTATGATGAAGATACAAATTCGACTATCCATT 540
 QY 541 ATTGCACCAACCAAGTAGTAAATCAAGTTAGTAAATGCTATTTTAGTGAAGATCCG 600
 DB 541 ATTGCACCAACCAAGTAGTAAATCAAGTTAGTAAATGCTATTTTAGTGAAGATCCG 600
 QY 601 GTTATCGTTTCTTAATGTCAGGACCAACGAGAGACGCTATTGATACAGATATAGTTAT 660
 DB 601 GTTATCGTTTCTTAATGTCAGGACCAACGAGAGACGCTATTGATACAGATATAGTTAT 660
 QY 661 GATGACCAAGATTATGTTTAAATCGATCTGCTGGTATGCTAAAAAGGAAAGATATAT 720
 DB 661 GATGACCAAGATTATGTTTAAATCGATCTGCTGGTATGCTAAAAAGGAAAGATATAT 720
 QY 721 GAATCACTGAGAAATATTCAGTATTAGAGCTTTAAAGCGATTGAACGTTCAAAATGTT 780
 DB 721 GAATCACTGAGAAATATTCAGTATTAGAGCTTTAAAGCGATTGAACGTTCAAAATGTT 780
 QY 781 GTTTTAGTGGTTATTGATGCAAGCAAGGATTTATGACCAAGATTAACGTTGTCAGGA 840
 DB 781 GTTTTAGTGGTTATTGATGCAAGCAAGGATTTATGACCAAGATTAACGTTGTCAGGA 840
 QY 841 TATGCATATGAACAGGTAAAGCAGTCTGATTGCTGCTAAATTAATGGGATACTGTGGA 900
 DB 841 TATGCATATGAACAGGTAAAGCAGTCTGATTGCTGCTAAATTAATGGGATACTGTGGA 900
 QY 901 AAAGATAGTAAACAGGATGAAGAAATTTGAAGATGAAGTACGTAAAGAAATTTCAATTTTA 960
 DB 901 AAAGATAGTAAACAGGATGAAGAAATTTGAAGATGAAGTACGTAAAGAAATTTCAATTTTA 960
 QY 961 GATTATGCACAAATTTGCTTTGCTGCTAAAGAGCGCAAGATTACGTATATTC 1020
 DB 961 GATTATGCACAAATTTGCTTTGCTGCTAAAGAGCGCAAGATTACGTATATTC 1020
 QY 1021 CCTTACATTAATGAAGCAAGTGAAGCAATTAAGAGCTGTTCAAAAGTTCAACTTTAAAT 1080
 DB 1021 CCTTACATTAATGAAGCAAGTGAAGCAATTAAGAGCTGTTCAAAAGTTCAACTTTAAAT 1080
 QY 1081 GAAGTCGTTACTGATGCAATTTCCATGCAACCTACACCAACAGGATAGAGGTTG 1140
 DB 1081 GAAGTCGTTACTGATGCAATTTCCATGCAACCTACACCAACAGGATAGAGGTTG 1140
 QY 1141 AATGCTTTTATGCAACCAAGTTGCTATAGAACCCAGCATTTGTTGTTTGTATTGTTAAAT 1200
 DB 1141 AATGCTTTTATGCAACCAAGTTGCTATAGAACCCAGCATTTGTTGTTTGTATTGTTAAAT 1200
 QY 1201 GATGTAGAAATTAATGTCATTTTCTTATTAACCGCTATTAGAGAAATCAAAATCCGCCCT 1260
 DB 1201 GATGTAGAAATTAATGTCATTTTCTTATTAACCGCTATTAGAGAAATCAAAATCCGCCCT 1260
 QY 1261 TTTGGCTTTTGAAGGTACACCAATTCATATTATAGTCTGGAAGAGA 1305
 DB 1261 TTTGGCTTTTGAAGGTACACCAATTCATATTATAGTCTGGAAGAGA 1305

RESULT 3
 AAF86461
 ID AAF86461 standard; DNA: 1311 BP.
 XX
 AC AAF86461;
 XX
 DT 26-JUN-2001 (first entry)
 XX
 DE Staphylococcus aureus yphC coding sequence.

XX	yphC; antimicrobial; cytostatic; antiulcer; microbial infection;
KW	gene therapy; vaccine; gastrointestinal carcinoma; gastric ulcer;
KW	gastritis; ds.
XX	Staphylococcus aureus.
OS	XX
FH	Location/Qualifiers
Key	1..1311
CDS	/tag- a
FT	/product= "Staphylococcus aureus yphC protein"
FT	
PN	WO200123418-A1.
XX	
PD	XX
XX	05-APR-2001.
PX	19-SEP-2000; 2000WO-US25566.
PF	XX
PR	28-SEP-1999; 99US-0406968.
PP	XX
PA	{SMIK } SMITHKLINE BEECHAM CORP.
PA	{SMIK } SMITHKLINE BEECHAM PLC.
PI	Zalacain M, Biswas S, Burnham MKR, Sylvester D, Mcdevitt D;
PI	Mathie TB;
DR	WPI: 2001-308138/32.
DR	P-PSDB; AA882089.
XX	
PT	Novel yphC polypeptides of Staphylococcus aureus useful for diagnosing
PT	and treating microbial infections, especially infection by
PT	Staphylococcus aureus and Helicobacter pylori -
PS	Claim 2; Page 2-3; 41pp; English.
XX	
CC	The present sequence is the gene encoding yphC polypeptide of
CC	Staphylococcus aureus. The yphC coding sequence and protein are useful
CC	for treating and diagnosing microbial infections such as infection caused
CC	by S.aureus and Helicobacter pylori. In addition, the yphC coding
CC	sequence and protein are useful for treating diseases such as
CC	H.pylori-induced cancers, e.g. gastrointestinal carcinoma, gastric
CC	ulcers, and gastritis. The present sequence was obtained from a library
CC	of clones of chromosomal DNA of S.aureus in E.coli. The sequencing data
CC	from two or more clones comprising overlapping S.aureus DNAs was used to
CC	construct the present contiguous DNA sequence.
SQ	Sequence 1311 BP; 451 A; 184 C; 278 G; 398 T; 0 other;
Query Match 99.0%; Score 1292.2; DB 22; Length 1311;	
Best Local Similarity 99.4%; Pred. No. 4.5e-287;	
Matches 1397; Conservative 0; Mismatches 8; Indels 0; Gaps 0;	
Oy	1 ATGACTAAACCTATAGTAGCTATTGTAGTAGGCCTAATGTAGTAAATCTACAATTTT 60
Dd	
Oy	1 ATGACTAAACCTATAGTAGCTATTGTAGTAGGCCTAATGTAGTAAATCTACAATTTT 60
Dd	
Oy	61 AATAAGAATAGTTGGAGAACGTTTTCGATGTTGAAGATACGCCAGTGTAACACGAT 120
Dd	
Oy	61 AATACAATAGTTGGAGAACGTTTTCGATGTTGAAGATACGCCAGTGTAACACGAT 120
Dd	
Oy	121 CGTATTTTATCTTCAGTGAGTGGTTAAACACATGATTTCAATATTATTGATACAGTGT 180
Dd	
Oy	121 CGTATTTTATCTTCAGTGAGTGGTTAAACACATGATTTCAATATTATTGATACAGTGT 180
Dd	
Oy	181 ATTGAATTTGCGATGACCACCATTCCAAACAAAATTAGACGCGAGCGAGAAATCGCCATA 240
Dd	
Oy	181 ATTGAATTTGCGATGACCACCATTCCAAACAAAATTAGACGCGAGCGAGAAATCGCCATA 240
Dd	
Oy	241 GATGAGCGGATGTTATTATTTTATGTTTAAAGTCCGCGTAAGGATTGACACAAAGCGAT 300
Dd	
Oy	241 GATGAGCGGATGTTATTATTTTATGTTTAAAGTCCGCGTAAGGATTGACACAAAGCGAT 300
Dd	
Oy	301 GAATGGTGCCTCAAAATTTTATACAAATCTAAAAAACCGGTCGTATTACCGGTTAACAAA 360
Dd	

XX	13-FEB-2002 (first entry)	
XX	Staphylococcus aureus DNA for cellular proliferation protein #1309.	
DE	Antisense; ds; prokaryotic cellular proliferation gene;	
XX	antibiotic; antibacterial; drug design.	
KW	Staphylococcus aureus.	
XX	WO200170955-A2.	
XX	27-SEP-2001.	
PD		
XX	21-MAR-2001; 2001WO-US09180.	
XX	21-MAR-2000; 2000US-191078P.	
PR	23-MAY-2000; 2000US-206848P.	
PR	26-MAY-2000; 2000US-207727P.	
PR	23-OCT-2000; 2000US-242578P.	
PR	27-NOV-2000; 2000US-253623P.	
PR	22-DEC-2000; 2000US-259331P.	
PR	16-FEB-2001; 2001US-269308P.	
XX	(ELIT-) ELITRA PHARM INC.	
PA	Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;	
PI	Yamamoto RT, Xu HH;	
PI	WPI: 2001-611495/70.	
DR	P-PSDB: AAU37138.	
DR	New polynucleotides for the identification and development of	
PT	antibiotics, comprise sequences of antisense nucleic acids -	
XX	ClaIm 27; Seq ID No 8634; SIlpp: English.	
XX	The invention relates to antisense inhibitors of genes essential to	
CC	prokaryotic cellular proliferation, their use in identifying the	
CC	genes, their use in the discovery of novel antibiotics, the essential	
CC	genes themselves and the encoded proteins. The prokaryotes used are	
CC	Escherichia coli, Staphylococcus aureus, Salmonella typhi, Klebsiella	
CC	pneumoniae, Pseudomonas aeruginosa and Enterococcus faecalis. The	
CC	invention is also useful for the identification of potential new targets	
CC	for antibiotic development. The antisense nucleic acids can also be used	
CC	to identify proteins used in proliferation, to express these proteins,	
CC	and to obtain antibodies capable of binding to the expressed proteins.	
CC	The proteins can be used to screen compounds in rational drug discovery	
CC	programmes. The antisense nucleic acid sequence is also useful to screen	
CC	for homologous nucleic acids which are required for cell proliferation in	
CC	a wide variety of organisms. The present sequence encodes an	
CC	essential prokaryotic cellular proliferation protein.	
CC	Note: The sequence data for this patent did not form part	
CC	of the printed specification, but was obtained in electronic	
CC	format directly from WIPO at	
CC	ftp.wipo.int/pub/published_pct_sequences.	
XX		
XX	Sequence 1311 BP: 452 A; 184 C; 278 G; 397 T; 0 other;	
QY	Query Match 98.9%; Score 1290.6; DB 23; Length 1311;	
DB	Best Local Similarity 99.3%; Pred. No. 1.1e-286;	
DB	Matches 1296; Conservative 0; Mismatches 9; Indels 0; Gaps 0	
QY	1 ATGACTAAACCTATAGTAGCTATTGTAGTAGGCGCTAATGTAGGTAAATCTACAATTTT 60	
DB	1 ATGACTAAACCTATAGTAGCTATTGTAGTAGGCGCTAATGTAGGTAAATCTACAATTTT 60	
QY	61 AATAGAATAGTTGGAGAACGGTGTTCGATTGTTGAAGATACGCCAGGTGTAAACACGAGAT 120	
DB	61 AATAGAATAGTTGGAGAACGGTGTTCGATTGTTGAAGATACGCCAGGTGTAAACACGAGAT 120	
QY	121 CGTATTATTTCTTCAGGTGAGTGGCTTAAACACATGATTTCATATTTATTCATACAGGTGCT 180	

OY 1261 TTGGTTTGAAGCTACACCAATTCATATTATAGCTCGAAAGAGA 1305
 Db |||||||

RESULT 5

AAV74669
 ID AAV74669 standard; DNA: 3621 BP.

XX AAV74669;

XX 16-MAR-1999 (first entry)

DE Staphylococcus aureus contig SEQ ID #358.

KW Computer readable medium; vaccine; S.aureus infection; immunodetection;
 KW cellulitis; eyelid infection; food poisoning; osteomyelitis; therapy;
 KW skin infection; surgical wound infection; scalded skin syndrome;
 KW toxic shock syndrome; ds.

OS Staphylococcus aureus.

XX Key Location/Qualifiers

PH misc_feature 481..540
 FT /*tag= a

FT /*note= "these bases represent a line of missing text in
 the sequence listing in the specification. They
 are included to maintain the nucleotide numbering
 given in the specification for this DNA sequence."

FT misc_feature 2281..2340

FT /*tag= b

FT /*note= "these bases represent a line of missing text in
 the sequence listing in the specification. They
 are included to maintain the nucleotide numbering
 given in the specification for this DNA sequence."

PN EP786519-A2.

XX 30-JUL-1997.

XX 07-JAN-1997; 97EP-0100117.

XX 05-JAN-1996; 96US-0009861.

XX (HUMA-) HUMAN GENOME SCI INC.

XX Barash SC, Choi GH, Dillon PJ, Fannon MR, Kunsch CA;
 PI Rosen CA;

XX WPI; 1997-374922/35.

XX Polynucleotide(s) and proteins derived from *Staphylococcus aureus*
 PT stored on computer readable medium and used in the production of
 PT anti-S.aureus vaccines

XX Claim 1: Page 1241-1243; 3271pp; English.

XX This sequence represents one of 5191 *Staphylococcus aureus* DNA sequences
 CC of the invention. The DNA sequences are recorded on a computer readable
 CC medium, preferably selected from a floppy or hard disk, random access
 CC memory (RAM), read-only memory (ROM) or CD-ROM. Homology searches using
 CC the S.aureus DNA sequences allows putative functions to be assigned so
 CC that protein-encoding or regulatory regions of commercial, therapeutic or
 CC industrial importance can be obtained. Specifically, sequences which are
 CC likely to encode antigens have been identified and these polypeptides can
 CC be used in a vaccine composition against *S.aureus* infection. The
 CC polypeptides can also be used in a kit for the immunodetection of
 CC S.aureus in a sample. S.aureus is implicated in numerous human diseases,
 CC including cellulitis, eyelid infections, food poisoning, osteomyelitis,
 CC skin and surgical wound infections, scalded skin syndrome, toxic shock
 CC syndrome, etc. Organisms transformed with the DNA sequences can be used
 CC for recombinant production of the polypeptides. The new DNA sequences

CC (and their fragments) are useful as primers or probes for isolating
 CC homologues of any of the *S.aureus* DNA sequences contained on the
 CC computer readable medium.

XX Sequence 3621 BP; 1279 A; 442 C; 692 G; 1085 T; 123 other;

Query Match 94.3%; Score 1230.2; DB 18; Length 3621;
 Best Local Similarity 94.6%; Pred. No. 1e-272;
 Matches 1235; Conservative 1; Mismatches 69; Indels 0; Gaps 0;

OY 1 ATGACTAAACCTATATAGTAGCTATTGTAGGTAGGCCTAATGTAGTAAATCTACAATTTT 60
 Db |||||||
 OY 2130 ATGACTAAACCTATATAGTAGCTATTGTAGGTAGGCCTAATGTAGTAAATCTACAATTTT 2189
 OY 61 AATAGAATAGTTGGAGAACGCTTTCCATTTGTTGAAGATACGCCAGGTGTACACAGAT 120
 Db |||||||
 OY 2190 AATAGAATAGTTGGAGAACGCTTTCCATTTGTTGAAGATACGCCAGGTGTACACAGAT 2249
 OY 121 CGTATTATTCTTCAGGTGAGTGTAAACACATGATTTCAATATTATTATGATACAGGTGT 180
 Db |||||||
 OY 2250 CGTATTATTCTTCAGGTGAATGCTTAACACNNNNNNNNNNNNNNNNNNNNNNNN 2309
 OY 181 ATTGAATTTGGTGTGCACCATTCCTCAACACAAATAGACGCCAGCAGCAAAATCCGCATA 240
 Db |||||||
 OY 2310 NNN 2369
 OY 241 GATGAAGCGGATGTTATTATTATTTTATGTTAACTGCGTGAAGGATTGACACAAAGCAT 300
 Db |||||||
 OY 2370 GATGAAGCGGATGTTATTATTATTTTATGTTAACTGCGTGAAGGATTGACACAAAGCAT 2429
 OY 301 GAAATGGTGGCTCAAAATTTTATACAAATCTAAACAAACCGGTGCTATTACGGTTAACAAA 360
 Db |||||||
 OY 2430 GAAATGGTGGCTCAAAATTTTATACAAATCTAAACAAACCGGTGCTATTACGGTTAACAAA 2489
 OY 361 GTAGATTAATATGGAATACCGGTACAGAGCTGTATGATTTCTATTTCATTAGGATTTGGTGA 420
 Db |||||||
 OY 2490 GTAGATTAATATGGAATACCGGTACAGAGCTGTATGATTTCTATTTCATTAGGATTTGGTGA 2549
 OY 421 CCGTATCCAATATACAGGTGCACATGTTTGGTCTTGGTGCATTTTGGTGCATTTTGGTGCATTTGTT 480
 Db |||||||
 OY 2550 CCGTATCCAATATACAGGTGCACATGTTTGGTCTTGGTGCATTTTGGTGCATTTTGGTGCATTTGTT 2609
 OY 481 TCTCATTTTGGTGAAGAGAGAGATCTTATGATGAAGATACAAATTCGACATATCCATT 540
 Db |||||||
 OY 2610 TCTCATTTTGGTGAAGAGAGAGATCTTATGATGAAGATACAAATTCGACATATCCATT 2669
 OY 541 ATTGGACGACCAACGCTAGGTAAATCAAGTTTGTAAATGCTATTTTAGGTGAAGATCGC 600
 Db |||||||
 OY 2670 ATTGGACGACCAACGCTAGGTAAATCAAGTTTGTAAATGCTATTTTAGGTGAAGATCGC 2729
 OY 601 GTTATCGTTTCTTAATGTTGCAGGCAACAGAGAGAGCTATTGATACAGAGTATATGTTAT 660
 Db |||||||
 OY 2730 GTTATCGTTTCTTAATGTTGCAGGCAACAGAGAGAGCTATTGATACAGAGTATATGTTAT 2789
 OY 661 GATGACAAAGATTTATGTTTAAATCGTATGCTGTTATGCTGTTAAAAAGGAAAAAGTATAT 720
 Db |||||||
 OY 2790 GATGACAAAGATTTATGTTTAAATCGTATGCTGTTATGCTGTTAAAAAGGAAAAAGTATAT 2849
 OY 721 GAATCAACTGAGAAATATTCAGTATTAAAGAGCTTTTAAAGAGCTTTTAAAGAGCTTTCAAAATGTT 780
 Db |||||||
 OY 2850 GAATCAACTGAGAAATATTCAGTATTAAAGAGCTTTTAAAGAGCTTTTAAAGAGCTTTCAAAATGTT 2909
 OY 781 GTTTTAGTGGTTATTGATGCGAACAAGGCAATTTTGAACAAGATACAAAGCTGTCGACGA 840
 Db |||||||
 OY 2910 GTTTTAGTGGTTATTGATGCGAACAAGGCAATTTTGAACAAGATACAAAGCTGTCGACGA 2969
 OY 841 TATGCACATGACAAAGGTAAAGAGCTGCTGATTGTCTGTAATAATGGGATACTCTGGAA 900
 Db |||||||
 OY 2970 TATGCACATGACAAAGGTAAAGAGCTGCTGATTGTCTGTAATAATGGGATACTCTGGAA 3029
 OY 901 AAGATAGTAAACAGTGAACAAATTTGAAGATGAAGTACGTAAAGAAATTTTCAATTTTCA 960
 Db |||||||
 OY 3030 AAGATAGTAAACAGTGAACAAATTTGAAGATGAAGTACGTAAAGAAATTTTCAATTTTCA 3089

QY 841 TATGACATGAACAGGTAAAGCAGTCTGATGATTCGTGAATTAATGCGATACACTGTGGA 900
 DB 51458 TATGCGCATGATCGCGAGCGGCTCATATTGTTAGTGAAACAATGGGATGCAATTAAC 51517
 QY 901 AAAGATAGTAAACGATGAAGAAATTTGAAGATGAAGTACGTAAAGAAATTCACATTTTA 960
 DB 51518 AAAGATGAAGAAACAATTAATGATGACAGAGATATTCGGGAGCAGTTCACATTCCTA 51577-
 QY 961 GATTATGACAAAATTTGCTGCTGCTAAAGAACGACAGATTAAGTACATATTC 1020
 DB 51578 AGCTATGACCAATTTGCTGATCTGCTAAACAACAACGCTTAACAACATTAATTC 51637
 QY 1021 CCTTACATTAATGAAGCAAGTGAACCACTAAACAGGTTCACAAAGTTCAACTTTAAAT 1080
 DB 51638 CGCGTCATTAACCAAGTAAGCGAATTAATTCATTAACGCTACAACTCTAGTATGCTAAAT 51697
 QY 1081 GAAGTCGTTACTGATGCAATTTCCATGAACCCCTACACACAGACAAAGGTAGAGCTTTG 1140
 DB 51698 GATGTTATTAGTATGCGGTTGCAATGAATCCATCAACCAATGATTAAGGTAAACGACTT 51757
 QY 1141 AATGCTTTTATGCAACACAGTTGCTATAGAACCCAGCCGACATTTGCTGATTTCTTAAT 1200
 DB 51758 AAATATTCTATACACACAGTGGCTGTTAAACCCGCAACATTTGCTGATTTGTTAAAT 51817
 QY 1201 GATGAGAAATTAATGCAATTTTCTTATTAACCGCTATTTAGAGAAATCAAAATCGGCCCT 1260
 DB 51818 GATCCAGAACTAATGCAATTTCTTATGAACGTTTCTTAGAAACCGGATTAGAGAGCA 51877
 QY 1261 TTTCGTTTGAAGTACACCAATTCATATTATAGCTCGAAAGAG 1304
 DB 51878 TTTCGTTTGAAGTACACCAATTCGAGTAATTTGCTCGTAAGCG 51921

RESULT 10

AB069245/c
 ID AB069245 standard; DNA; 3011208 BP.
 XX
 AC AB069245;
 XX
 DT 29-AUG-2002 (first entry)
 XX
 DE Listeria innocua DNA sequence #684.
 XX
 KW Antibacterial; Listeria; food contamination; mutational analysis;
 XX Infection; ds.
 XX
 OS Listeria innocua.
 XX
 PN W0200228891-A2.
 XX
 PD 11-APR-2002.
 XX
 PF 04-OCT-2001; 2001WO-FR03061.
 XX
 PR 04-OCT-2000; 2000FR-0012697.
 XX
 PA (INSP) INST PASTEUR.
 XX (CNRS) CNRS CENT NAT RECH SCI.
 XX
 PI Kunst F. Glaser P;
 XX
 DR WPI: 2002-332479/37.
 XX
 PT New genomic sequences from Listeria species, useful for detection,
 XX treatment and prevention of infection, also related polypeptides,
 XX antibodies and modulators .
 PS Claim 5; SEQ ID 2058; 180pp; French.
 XX
 CC The present invention relates to nucleic acid sequences
 CC (AB067188-AB071212) from Listeria sp. The sequences are useful as probes
 CC and primers for identification and/or detection of Listeria (e.g. as

CC contaminants in foods, or mutational analysis) and for analysis of
 CC gene expression. Proteins encoded by the nucleic acid sequences can be
 CC used to screen for compounds that modulate gene expression, replication
 CC and pathogenicity of Listeria (potential therapeutic agents), also for
 CC treating infections by Listeria, and are useful as immunogens in
 CC anti-Listeria vaccines.
 CC Note: The sequence data for this patent did not form part
 CC of the printed specification, but was obtained in electronic format
 CC directly from WIPO at [ftp.wipo.int/pub/published_pct_sequences](http://wipo.int/pub/published_pct_sequences).

XX
 SQ Sequence 3011208 BP; 941651 A; 568176 C; 559189 G; 942192 T; 0 other:

Query Match 51.9%; Score 676.8; DB 24; Length 3011208;
 Best Local Similarity 69.9%; Pred. No. 6.7e-145;
 Matches 912; Conservative 0; Mismatches 392; Indels 0; Caps 0;
 QY 1 ATGACTAAACCTATATAGTACTATTGCTAGGTAGGCTAATGTAGTAAATCTACAATTTTT 60
 DB 2066810 ATGGCAAAACCGATTGTAGCGATTGTGCGACGTCCAACCGTTGCAAAATCGCATTTT 2066751
 QY 61 AATAGAAATAGTTGGAGAACGTTGCTGATTTCTTCAAGATACGCCAGGTGTAAACACGAT 120
 DB 2066750 ACAGAAATCGTTGGTGAACGTTGTTCCATAGTGAAGATGTTCCCGGTGTGACACGTGAC 2066691
 QY 121 CGTATTATTTCTTCAAGTGAAGTGTAAACATGATTTCAATATTTATGATACAGTGGT 180
 DB 2066690 CGCATATATAATTCAGCGGAATGGCTTGGAAAGAAATTTAAACATTTATTGATACAGTGGT 2066631
 QY 181 ATTGAATAGTGTGATCGACCATTCACAAACAAATATAGCGGACGAGCAAAATCGCCATA 240
 DB 2066630 ATGATCTTTCCGACGACCATTTCTAGAGCAAAATTCGCCACACGCGAAATCGCAAT 2066571
 QY 241 GATGAAGCGGATGTTATTATTTTATGGTTAACTGCGTGAAGGATTGACACAAACGAT 300
 DB 2066570 GATGAAGCAGACGTAATTTATTTTATACCAATGTCGTGAAGGGTTACCGATGACAC 2066511
 QY 301 GAAATGGTCGCTCAAAATTTTATACAAATCTAAACAAACCGTCGCTATTAGCGGTTAAACAA 360
 DB 2066510 GAACAGTAGCAAAATTTCTTTACGGTCTAATAAACCAATTTGTTTACGATTAATAA 2066451
 QY 361 GTAGATAATATGGAATGCGTACAGAGCTGTATGATTTCTTATTAGTATGATTTGGTGA 420
 DB 2066450 GTAGATAACCCAGAAATGCGTGATCAGATTTATGATTTTCTTCTTGGGTTGGTGAG 2066391
 QY 421 CGGTATCCAATATCAGGCTCAGATGTTAGTCTTGGTGAAGTGTGATGATGATGATGATGAT 480
 DB 2066390 CGGTATCCAATTTCTGTTCTCATGGACTAGGCTTGGTGAAGTGTGATGATGATGATGAT 2066331
 QY 481 TCTCATTTTGGTGAAGAGGAAGATCCTTTATGATGAAGATACAAATTCGACTATCCATT 540
 DB 2066330 GCTCATTTTCCAAGAGAGAGAGGAAGATACCCAGATCAACAGTGAATTTAGTTTG 2066271
 QY 541 ATTGACGACCAACGATAGGTAAATCAAGTTTAGTAAATGCTATTTTAGTGAAGATCCG 600
 DB 2066270 ATTGTCGCGCAAAATTTGGTAAATCAATCTATTCTAAACGACCTTCTTGGCGAAGACC 2066211
 QY 601 GTTATCGTTTCTTAATGTTGCGAGGACACGAGACGCTATTGATACAGATATAGTTAT 660
 DB 2066210 GTCATTTGTTCTTGACATTTCCGGAACCTACTCGTGATGCAATTCATACAACTTT 2066151
 QY 661 GATGACAGATTTATGTTTAAATCGACTACTGCTGATGCTGCTGCTGCTGCTGCTGCTGCT 720
 DB 2066150 GATGTCAGGATTTATGTCATGTTGATACCGGAATGAGAAACGCTGGGAAGTGTAT 2066091
 QY 721 GAATCACTGAGAAATATTTCAATATTAAGAGCTTTTAAACCGGATTAAGAGCTTTCAATGTT 780
 DB 2066090 GAAGCAGAGAGAAATATAGTGTGTTTACGTCATGAGACAAATTAAGACGCTCCGATGTT 2066031
 QY 781 GTTTTAGTGTATTGATGACAGAAAGGCAATTTTGAACAAAGATTAACAGCTGTTGTCAGCA 840
 DB 2066030 GTTCTGTGTTATCAACGACAGAGAGGATTTCTGTCAGCAAGATTAAGCGGATTCGTCGA 2065971
 QY 841 TATGACATGAACAGGTAAAGCAGTCTGCTGATTTGCTGATTAATTAATGCGGATCTGTGGA 900

[illegible]

RESULT 11

AAH53683	ID	AAH53683	standard; DNA; 960 BP.
XX			
XX			
AC			
XX		AAH53683;	
XX			
DT		03-SEP-2001	(first entry)
XX			
DE		S. epidermidis	open reading frame nucleotide sequence SEQ ID NO:2759.
XX			
XX		Staphylococcus epidermidis	SRI strain; infection; diagnosis;
KW		vaccination; endocarditis; ds.	
KW			
OS		Staphylococcus epidermidis.	
XX			
PN		WO200134809-A2.	
XX			
PD		17-MAY-2001.	
XX			
PF		09-NOV-2000; 2000WO-US30782.	
XX			
PR		09-NOV-1999; 99US-0164258.	
XX			
PA		(GLAX) GLAXO GROUP LTD.	
XX			
PI		Klimmerly WJ;	
XX			
DR		WPI; 2001-316495/33.	
XX		P-PSDB; AAG82833.	
XX			
PT		Nucleic acids encoding polypeptides	from Staphylococcus epidermidis,
PT		useful for vaccinating against infections,	e.g. endocarditis .
XX			
PS		Claim 8; Page 718-719; 2188pp;	English.
XX			
CC		AAH52304 to AAH53970	represent nucleic acids (I) encoding polypeptides
CC		(II), given in AAG81454 to AAG83120,	from Staphylococcus epidermidis.
CC		(i) and (ii) can have antibacterial activity	and therefore can be used
CC		in vaccination. The nucleic acids (I) may	be used to produce the
CC		S. epidermidis polypeptides (II) via the	production of vectors
CC		containing them which are used to produce	hosts cells which express the

1589796 CAATCATCTGTCAGTCAGTGGAGCTGATATAGTGTAACTTCCAAATCTTGTCTTA 1589797
 966 TGCACAAATTCGTTTGTGCTGCTAAAGACGCAAGATTACGTATATTCCTTA 1025
 1589736 TGCACAAATTCGTTTGTGCTGCTAAAGACGCAAGATTACGTATATTCCTTA 1589677
 1026 CATTAAATGAAGCAAGTGAACCATATAAAGCTGTTCAAAGTTCACATTTAAATGAAT 1085
 1589676 GATTAAACGTATTAGCGCAAGTCAAAACCAACGATTTCGCTGCTGTATTAATGATG 1589617
 1086 CGTTACTGATGCAATTCCTGATGACCTACACACACACAGTGTGATGAT 1145
 1589616 GATTATGATGCTATTCCTATTAATCCACACACACCTGATTAAGGAAACGCTTAAGAT 1589557
 1146 CTTTATGCAACACAGTGTGCTATAGACACCGCATTTGTTGTTTAAATGATG 1205
 1589556 TTTCTAGCGCAACTCAAGTTGCTGTTAAACCAACCACTTTCGTTGTTTAAACGAAGA 1589497
 1206 AGAATTAATGCAATTTCTTATTAACGCTATTAGAGAAATCAAAATCCGTCGCTTTGG 1265
 1589496 AGAATGATGCACTTTTCATATCTTCTTCTTACAAATCAAAATCCGTCGATTTGT 1589437
 1266 TTTTGAAGGTACACCAATTCATATATAGCTCGAAA 1301
 1589436 TTTTGAAGGAACCAACCAATTAATCTAATAGCAGTAA 1589401

RESULT 14
 ABK75008
 ID ABK75008 standard; DNA: 1311 BP.
 AC ABK75008;
 DT 13-AUG-2002 (first entry)
 DE Bacillus licheniformis genomic sequence tag (GST) #2299.

KW Differential gene expression; genomic sequenced tag; GST;
 KW altered culture condition; environmental stress;
 KW physiological provocation; ds.

OS Bacillus licheniformis.

PN WO200229113-A2.

XX 11-APR-2002.

XX 05-OCT-2001; 2001WO-US31437.

XX 06-OCT-2000; 2000US-0680598.

PR 27-MAR-2001; 2001US-279526P.

XX (NOVO) NOVOZYMES BIOTECH INC.

PA (NOVO) NOVOZYMES AS.

XX Berka R, Clausen IG;

XX WPI: 2002-416684/44.

Monitoring differential expression of several genes in first Bacillus
 cell relative to expression of same genes in one or more second
 Bacillus cells, by using substrate containing Bacillus genomic
 sequenced tag array

PS Claim 4: SEQ ID NO 2299; 200pp; English.

The invention describes a method of monitoring differential expression of
 genes in a first Bacillus cell relative to expression of the genes in
 other Bacillus cells, comprising hybridising labelled nucleic acid probes
 isolated from Bacillus cells to a substrate containing array of Bacillus
 genomic sequenced tags (GST), examining the array, and determining
 relative gene expression by an observed hybridisation reporter signal of

a spot in the array. The method is useful for measuring the expression of
 genes in a first Bacillus cell relative to expression of the same genes
 in one or more second Bacillus cells. The method is useful for monitoring
 global expression of several genes from a Bacillus cell, discovering new
 genes, identifying possible functions of unknown open reading frames and
 monitoring gene copy number variation and stability. Monitoring changes
 in expression of genes may be used to provide a representation of the way
 in which Bacillus cells adapt to changes in culture conditions,
 environmental stress or other physiological provocation. Extensive
 follow-up characterisation is unnecessary, when one spot on an array
 equals one gene or one open reading frame, since sequence information is
 available. This sequence represents a genomic sequence tag (GST) used in
 the method of the invention.
 Note: The sequence data for this patent did not form part of the printed
 specification, but was obtained in electronic format directly from WIPO
 at
 ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 1311 BP; 376 A; 279 C; 353 G; 303 T; 0 other;

Query Match 45.98; Score 599.4; DB 24; Length 1311;
 Best Local Similarity 66.2%; Pred. No. 5.3e-128;
 Matches 864; Conservative 0; Mismatches 441; Indels 0; Gaps 0;

QY 1 ATGACTAAACCTATAGTAGCTATTAGTAGTAGCCCTAATAGTAGTAAATCTACAATTTT 60
 DB 1 ATGGGAAAACCTGCTGCTAGCCATAGTTGGAAGACCCCAATGTGGGAAGTCCACGATCTT 60
 QY 61 AATAGAATAGTTGGAGAACCTGTTTCGATTGTTGAAGATACGCCAGTGTAAACAGAGAT 120
 DB 61 AACAGGATCGCGGCGGAAGAAATTTTCGATTGGAAGATACACCTGCGGTGACCGGGAT 120
 QY 121 CGTATTATTCTTCAGGTGAGTGGTTAAACACATGATTTCATATTATGATACAGGTGGT 180
 DB 121 CGGATATACAGCTCTGCGGAGTGGCTGAATACAGACTTCAACGCTGATGATACAGGGGA 180
 QY 181 ATTGAATTTGGTGATGCCACCATTTCCAAACACAAATFAGAGCGCAGCGAAATGCCATA 240
 DB 181 ATCGAAGTCGCGGATGAGCGTTTTCGCGCAGATTCGCGCATCAGCGCGAGATGCCATG 240
 QY 241 GATGAAGCGGATGTTATTATTTTATGTTTACGTCGTTGAAGGATTCACACAAAGCGAT 300
 DB 241 GAAGACGAGATGATTATTTTCATGACGACGCGCGGAGGCGTCCAGCCGCTGAT 300
 QY 301 GAAATGTCGCTCAAAATTTTATACAAATCTAAACAAACCGTGTATTAACGGTTAAACAAA 360
 DB 301 GARGAAGTCGCGGAAAATTTTATACCGCTACAAAAAAACCCGCTGCTTCGCGGTTTAAATA 360
 QY 361 GTAGATAATATGGAATTCGCTACAGACGCTGTATGTTCTATTATTAGGATTTGGTGAA 420
 DB 361 GTGGATAATCTGGAATGAGACCAACATTTACGATTTTATGCCCTCGGCTTTGGCGAG 420
 QY 421 CCGTATCCAATATCAGGTCACATGGTTTAGGTCTTTGGTGACTTGTAGATGCAATGTTGT 480
 DB 421 CCGTTCCCGATTTTCATGGACGCTATGCTTGTGCGGATGTCGCTTTCAGCCGCTGAGC 480
 QY 481 TCTCATTTTGTGGAAGAGGAAGATCTCTTATGATGAAGATACAAATTCGACTATCCATT 540
 DB 481 GAACATTTTAAAAACATACCCGAGACGAAGTATGAGGATGAGTCTGTTCAATTTTGCCTC 540
 QY 541 ATTGGACGACCAACAGCTAGTAAATCAAGTTTAGTAAATGCTATTTTAGGTGTAAGATCGC 600
 DB 541 ATCGGCGCGCGCAACGTCGGAAAATCTCTCCCTTGTAAACGCGCATGATTTGGAGAGAGCGC 600
 QY 601 GTTATCGTTTCTTAATGTTGCGAGGACACGAGACGCTATTGATFACAGAGTATAGTTAT 660
 DB 601 GTTATCGTCAGCAATATTGCGGGAACGACCGCGACGCTTTGACACGAGGTTTACATAC 660
 QY 661 GATGCAACAGATGTTTAAATCGATCTGCTGATTCGCTGTAACGCTGTAACGCTGTAACGCT 720
 DB 661 AATCAAGACACTTTGTCATCGTGGATACAGCGCGGATGAGAAAAGGAAAGTATGTT 720
 QY 721 GAATCAACTGAGAAATATTTCAGTATTAAAGAGCTTTAAAGCGATTGAAAGCTTCAAAATGTT 780

Db 721 GAGGCACTGAAATAACAGCGTCTCCGAGCTTTAAAGCGGATCGACCGTTTCAGAGTC 780
 Qy 781 GTTTTAGTGGTTATTGATGCAACAGGCAATATTGMAACAAGATAAAGCTGTTCAGGA 840
 Db 781 GTCTCGCGTCTTGAGCGCGGAAGAAGGCAATATTGAGCAGGACAAAGCGGATTCGCGG 840
 Qy 841 TATGCACATGAACAGGTAAGCAGTCTGATTTGCTGTAATAAATGGGATAGTGGAA 900
 Db 841 TATGCCATGAAGCAGGAAAGCTGTCTCATCGTCTGTAACAATAATGGGATCGAGTCGAA 900
 Qy 901 AAAGATAGTAAACAGCAGCAAGAAATTTGAAGATGAAGTACGTAAAGAAATTCCAATTTTA 960
 Db 901 AAAGACGAGCGGACGATGAAGAAATTTGAAGAAAGGTTAGAGACCACTTTCATTTTGG 960
 Qy 961 GATTATGACAAATTTGCTTTGCTGCTTAAGAACGCAAGATTTAGCTACATATTTC 1020
 Db 961 GACTACGCGCTGTGTTGTTTATGTCGCTTGGCTTAAGAAAGCGCATCCATACGCTGATG 1020
 Qy 1021 CCTTACATTAATGAAGCAAGTGAACCAATTAACCAATTAACCAATTAACCAATTAACCAAT 1080
 Db 1021 CCTCGGTCATGACGCGGAGCAACCAATTCGATGAGGTCGAGACCAATATTCCTTAAT 1080
 Qy 1081 GAAGTCGTTACTGATGCAATTTCCATGACCCCTACACCAAGCAAGCAAGTACAGCTTG 1140
 Db 1081 GATATCATATTGATGCGGTCGCTGATGACCCGCGGCGGCTAATGGAACCGGTTA 1140
 Qy 1141 AATGCTTTTATGCAACAAAGTTGCTATAGAACCCAGCAATTTGCTGATTTGTTAAT 1200
 Db 1141 AAAATTTATATGCGACAAAGTGGCTGTGAAGCCGCGGACATTTGCTGATTTGTCAT 1200
 Qy 1201 GATGATGAATTAATGCATTTTCTTATTAACGCTATTATAGAGAAATCAAAATCCGTCGCT 1260
 Db 1201 GACCCTGAGCTGATGCAATTTTCTTACGAAAGCTTTTGGAAACCGGATACGCGAGGCC 1260
 Qy 1261 TTTGCTTTGAGGTACACCAATTCATTTATAGCTCGGAAGAGA 1305
 Db 1261 TTCGATTTGAAGGAACCACTCAGGATTTTCGANGAGCAAGA 1305

RESULT 15

ID ABA90521
 XX ABA90521 standard; DNA; 2365589 BP.

AC ABA90521:

XX 16-MAY-2002 (first entry)

XX Genomic sequence of *Lactococcus lactis* IL1403.

DE Biosynthesis; biodegradation; lactic bacterium; yogurt; cheese; ds.

KW *Lactococcus lactis* IL1403.

OS FR2807446-A1.

XX 12-OCT-2001.

XX 11-APR-2000; 2000PR-0004630.

XX 11-APR-2000; 2000PR-0004630.

XX (INRG) INRA INST NAT RECH AGRONOMIQUE.

PA Bolotline A, Sorokline A, Renault P, Ehrlich SD;

PI wPI; 2002-043418/06.

XX New nucleotide sequence useful in the identification of *Lactococcus*

PT *lactis* and related species -

XX Claim 1; SEQ ID 1; 2504pp; French.

CC The present invention is related to a *Lactococcus lactis* nucleotide
 CC sequence (ABA90521) and related proteins (ABBS3300-ABBS5621). The
 CC nucleic acid sequence is useful in the detection and/or amplification of
 CC related species. The proteins of the invention are useful for the
 CC biosynthesis or biodegradation of a composition of interest. The
 CC invention helps research in lactic bacteria, particularly useful in the
 CC production of yogurt and cheese.
 CC Note: The sequence data for this patent is based on equivalent patent
 CC W200177334 (published 18-OCT-2001) which is available in electronic
 CC format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.
 XX
 SQ Sequence 2365589 BP; 765914 A; 415261 C; 420487 G; 763977 T; 0 other;

Query Match 44.6%; Score 582.4; DB 24; Length 2365589;

Best Local Similarity 66.5%; Pred. No. 3e-123;

Matches 867; Conservative 0; Mismatches 431; Indels 6; Gaps 2;

Qy 1 ATGACTAAACCTATAGTAGCTATTGTAGTAGGCTAATGTAGTAAATCTACAATTTTT 60
 Db 761403 ATGAGCTTACCTACAGTAGCCATCGTTGGCGTCCAAATGTCGGAATAACAAGATATTC 761462
 Qy 61 AATAGAAATAGTTGGAGAACGCTGTTTCGATTCTCAAGATACGCCAGCTGAACACGAGAT 120
 Db 761463 AACCGTATTGACGAGAACGCAATCTCAATTGTCGAAGACATTCGAGGTGTAACCTGAC 761522
 Qy 121 CGTATTATTCTTCAGGTGAGTGGTTAAACACATGATTTCAATATTATTGATACAGGTGT 180
 Db 761523 CGTATCTATGCCACAGGAGAAATGTTAAACCGTAAATTTCAATATCATTCGATACAGGTGG 761582
 Qy 181 ATTGAATTTGGTGTATGTCACCATTTCCAAACACAAATTAGAGCGGAGGACAGAAATCGGCATA 240
 Db 761583 ATTGAATTTTCAGACGAACCTTTTCATGATCTGAATTCGAGCGCAAGCTGAATTCGATG 761642
 Qy 241 GATGAGCGGATGCTTATTATTATTTTATGCTTAACCTGCGTGAAGGATTGACAAACGCGAT 300
 Db 761643 ACCGAACGAGATGTTATCATCGCTGTTGTCGCGGAGAACAGGTATTACAGATCCGAT 761702
 Qy 301 GAATGTTGCTCAAAATTTTATACAAATCTAAAAACCGGTCGTATTAGCGGTTAACAAA 360
 Db 761703 GAAGCGTGTCTTAATTAATTTATGCTACAGATAAACCCAGTTATCTGTTGCTCAATAA 761762
 Qy 361 GTAGATAATATGAAATCGCTACAGACGTGTATGATTCTTATTCTATTAGGATTTGTTGAA 420
 Db 761763 GTTGATAACCCAGAACGTCGAATCGAATCTTCTACTACTTGGACTTGGCGAT 761822
 Qy 421 CGTATCCAAATATCAGGTCACATGCTTGTAGGCTCTTGTCTGCTGTTGTTAGACTGTT 480
 Db 761823 CTTATCTGTTTACGCGGTTTACGCGGATTTGGGACAGGATGTCCTTGTGCAATCGTT 761882
 Qy 481 TCTCATTTTGTGAAGAGAAAGATCTTTATGATGAGATACAATTCGACTATCCATT 540
 Db 761883 CAATAATCTTCTTAATGAATTTGAAGA---AGAAAATGAGAAATGTCATCAAGTTCAGCTTG 761939
 Qy 541 ATTGGACGACCAACAGCTAGGTAAATCAAGTTTGTAGTAAATGCTATTATTAGGTGAAGATCC 600
 Db 761940 ATTGGCGCCCAACAGCTTGGGAAATCATCTCTTATTAACTGATTTCTTGGCGAAGACCG 761999
 Qy 601 GTTATCTTCTTAATGTTGCGAGGACAAAGACACCTATTGATACAGAGTATAG---T 657
 Db 762000 GTAAATGCAAGTCCAAATTCGAGGACAACTCGTATGCGGATGACACACACATTTGTCGAT 762059
 Qy 658 TATGATGCAAGATTTATGTTTTTAATCGATACTGCTGTTATGCTGTAAGGATAAGGAAAGTA 717
 Db 762060 TCTGAAGACCAAGAATTTGTTATGATGATACCCCGGAATGCGTAAATCAGGAATAAT 762119
 Qy 718 TATGAATCACTGACAAATATTTCAGATTATAGAGCTTTTAAACGCTTCAAGCTTCAAT 777
 Db 762120 TACGAATAATCTGAAAAAATCTCAGTCATGCGTCCATGCGTCCATGCGCTTCAGAT 762179
 Qy 778 GTTGTGTTTATGTTTATTGATGCAAGAAAGGCAATTTATGCAACAAGATAAAGCTGTGCA 837
 Db 762180 ATCGTTTTGATGTTGATTAAGCTTGAAGGAAATTCGGAATATGATGCTGCTCGCT 762239

OY 838 GGATATGCACATGAACAGGTAAAGCAGTCGTGATTCGTGAATAAATGGGATACGTGTG 897
Db 762240 GGATTTGCTCAGCAGCTGGTAAAGGAATTTTGATGTGTGTCATTAATGGGATACCCTT 762299
OY 898 GAAAAGATAGTAAACAGATGAAGAAATTTGAAGATGAAGTACGTAAAGAAATTCCAAATTT 957
Db 762300 GAGAAAGATAACGATACAAATGAAAACCTTTGAATTAGAAATTCGGACAAAGTTCAAATTC 762359
OY 958 TTAGATTATGCACAAATTCCTTTGTGCTGCTAAAGACGCACAAGATTACGTACATTA 1017
Db 762360 CTTGATATGCACCAATCTTTATGTATCTGCTAAGACAGACACGCTCTGAATAAACTC 762419
OY 1018 TTCCCTTACATTAATGAAGCAAGTGAAGAACCAATAAAAAACGTGTTCAAAAGTTCAACTTTA 1077
Db 762420 CCAGATATGATTAAAGAAATTCATCATGACAAAACCTTACGTATTTCAAGTTCTGTTTTG 762479
OY 1078 AATGAAGTCGTTACTGATGCAATTTCCATGAACCCCTACACCAACAGACAAAAGGTAGACGT 1137
Db 762480 AATGATGTCATTTATGATGCTGTAGCGATTATCCAACTCCGACGGACAAAGGAAAACGT 762539
OY 1138 TTCAATGCTCTTTTATGCAACACAAAGTTGCTATAGAACCCACGACATTTGTTGTTTGT 1197
Db 762540 TTGAAGATTTTCTATGCAACACAAAGTAGCGATTAAACCCAACTTTTGTGTTTGTGTC 762599
OY 1198 AATGATGTAGAATTAATGCAATTTTCTTATAACGCTATTTAGAGAATCAAAATCCGTGCC 1257
Db 762600 AATGAAGAAGAACTCATGCACCTCTCATATCTTCGGTTCTTGAANAATCAAAATCCGTAAA 762659
OY 1258 GCTTTTGGTTTTGAAGGTACACCAATTCATATTATAGCTCGAAA 1301
Db 762660 GCCTTGTGTTTTGAGGAAACACGATTCATTTGATTCACGTA 762703

Search completed: September 12, 2003, 17:32:19
Job time : 448.085 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 12, 2003, 17:17:59 ; Search time 3513.16 Seconds
(without alignments)
9028.139 Million cell updates/sec

Title: US-09-815-242-4228

Perfect score: 1305

Sequence: 1 agactaacctatagtagc.....atattatagctcgaagaga 1305

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

EST:*

- 1: em_estba:*
- 2: em_esthum:*
- 3: em_estlin:*
- 4: em_estmu:*
- 5: em_estov:*
- 6: em_estpl:*
- 7: em_estro:*
- 8: em_htc:*
- 9: gb_est1:*
- 10: gb_est2:*
- 11: gb_hic:*
- 12: gb_est3:*
- 13: gb_est4:*
- 14: gb_est5:*
- 15: em_estfun:*
- 16: em_estom:*
- 17: em_gss_hum:*
- 18: em_gss_inv:*
- 19: em_gss_pln:*
- 20: em_gss_vrt:*
- 21: em_gss_fun:*
- 22: em_gss_nam:*
- 23: em_gss_mus:*
- 24: em_gss_pro:*
- 25: em_gss_rtd:*
- 26: em_gss_phg:*
- 27: em_gss_vrl:*
- 28: gb_gss1:*
- 29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	486.6	37.3	2662	28	BH770770 LLMtag51
2	143.6	11.0	552	9	AI390536 mu83b02.y
3	128.4	9.8	721	12	BJ573090 BJ573090
4	124	9.5	1233	11	AY110834 Zea mays

5	118.6	9.1	583	14	CA782602
6	108.4	8.3	431	9	AA215138
7	99.8	7.6	579	13	BQ514303
8	99.2	7.6	532	13	BQ281519
9	89.4	6.9	395	9	AI465410
10	88.4	6.8	745	14	CD377507
11	84.6	6.5	171	28	BH234869
12	83.2	6.4	1101	29	CNS0039G
13	81.4	6.2	167	28	BH234795
14	79.6	6.1	653	13	BQ855888
15	78.8	6.0	501	14	N37967
16	78.8	6.0	613	10	BG293568
17	77.4	5.9	534	14	CA938405
18	76.8	5.9	631	28	BH661059
19	74.2	5.7	908	28	CD407144
20	74.2	5.7	908	28	AZ548467
21	73.2	5.6	1359	28	BH770563
22	73	5.6	769	13	BQ802107
23	72.6	5.6	1010	14	CD102897
24	71.6	5.5	410	14	CD474342
25	71.6	5.5	841	14	CB996761
26	71.6	5.5	905	28	AZ550256
27	71.4	5.3	487	12	BI272296
28	71	5.4	538	13	BQ490035
29	70.6	5.4	552	14	CA742475
30	70.6	5.4	623	12	BJ463635
31	69.8	5.3	843	28	AZ551618
32	69.8	5.3	931	28	BH160372
33	69.2	5.3	412	12	BM402985
34	69	5.3	772	12	BI920601
35	68.8	5.3	849	28	AZ546009
36	68.4	5.2	912	28	AZ551092
37	67.2	5.1	745	28	AY079759
38	66	5.1	911	28	BH164417
39	65.8	5.0	512	10	AW979775
40	65.6	5.0	541	13	BQ627620
41	65.6	5.0	541	13	BQ627620
42	65.2	5.0	571	13	BU578174
43	64.2	4.9	721	10	BG646649
44	64	4.9	350	14	CA735279
45	63.6	4.9	356	14	H37009

ALIGNMENTS

RESULT 1
BH770770
LOCUS BH770770 2662 bp DNA linear GSS 01-MAY-2002
DEFINITION LLMtag515 MG1363 Random Sequence Tag Library Lactococcus lactis
subsp. cremoris genomic, genomic survey sequence.
ACCESSION BH770770
VERSION BH770770.1 GI:20373727
KEYWORDS GSS.
SOURCE Lactococcus lactis subsp. cremoris
ORGANISM Lactococcus lactis subsp. cremoris
Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
Lactococcus.

REFERENCE 1 (bases 1 to 2662)
AUTHORS Bolotin, A., Ehrlich, S. D. and Sorokin, A.
TITLE Studies of genomes of dairy bacteria Lactococcus lactis
JOURNAL Sci. Alimentis, (2002) In press
COMMENT Contact: Sorokin A
Genetique Microbienne
INRA

CRJ INRA, Domaine de Vilvert, 78352 Jouy en Josas cedex, France
Tel: 33 1 34 65 25 16
Fax: 33 1 34 65 25 21
Email: sorokine@jouy.inra.fr
best homologue in strain IL1403 is yphL (98%)
Class: Shotgun
High quality sequence start: 30
High quality sequence stop: 2634.

FEATURES		Location/Qualifiers	
source	1..2662	/organism="Lactococcus lactis subsp. cremoris"	
		/mol_type="genomic DNA"	
		/strain="MGI363"	
		/db_xref="taxon:1359"	
		/clone_lib="MGI363 Random Sequence Tag Library"	
		/note="Vector: pSGM2; Site_1: SmaI; Library of chromosomal fragments of L.lactis strain MGI363 was prepared by partial AluI digestion or by sonication."	
BASE COUNT	871 a 506 c 513 g 772 t		
ORIGIN			
Query Match	37.3%; Score 486.6; DB 28; Length 2662;		
Best Local Similarity	65.6%; Pred. No. 5.2e-99;		
Matches 743; Conservative 0; Mismatches 384; Indels 6; Gaps 2;			
QY	1 ATGACTAAACCTATAGTAGCTATTGTAGGTAGGCGCTAAATGTAGTAAATCTACAAATTTT 60		
DB	1533 ATGAGCTTACCTACAGTCGCCATCGTTGGCGCTCCAAATGTTGGAAATATCAAGATATTC 1592		
QY	61 AATAGATAGTGGAGACGCTGTTTCGATTGTTGAAGATACGCCAGGTGTAAACAGCAT 120		
DB	1593 AACCGTATGCGAGGAGACGATCTCAATGTGCGAAGACATTCACAGGTGTAACCTGTCAC 1652		
QY	121 CATTATTTATCTTACGCTAGCTGTTAAACACATGATTTTCAATATTTGATACAGGTGTT 180		
DB	1653 CGTATCTATGCCACAGGAGATGTTAAACCGCTAAATTCACATCATCGACACCGGTGG 1712		
QY	181 ATTGAATTTGGTATGACCACTTCCAAACACAAATTAGACGCGAGCAGCAAAATCCGCATA 240		
DB	1713 ATTGAATTTTCAGATGAACCTTTTCATGACTGAAATTCGGCGACAGCGCAAAATCCGCATG 1772		
QY	241 GATGAAGCGGATCTATTATTTTATGTTAACTGCTCGTGAAGGATGTGACAAACGCGAT 300		
DB	1773 ACCGAGCAGATGTTATTCGCTGCTGTTGATGGAGAACAGCAATTAATGATGCCAT 1832		
QY	301 GAATGTCGCTCAAAATTTATACAAATCTAAACACCGGCTGCTATAGCGGTTAAACAAA 360		
DB	1833 GAAGCAGTGTGTAATTTCTTATTCGTACAGATTAACACCTGTCATTTATCGTCAATAAA 1892		
QY	361 GTAGATAATATGAAATCGCTACAGACGCTGTATGATTTCTATTCATTTAGGATTTGGTGA 420		
DB	1893 GTCGAAACCCGGAAGCTGGAATTTTGTGACTTCTATTCACCTGCGACGCGCAT 1952		
QY	421 CCGTATCCATATCAGGTCACATGCTTTAGGCTTGGTGACTTGTATGATCAGTTGTT 480		
DB	1953 CCATATCTGTATGCTGTTTCACGGGATTTGGACAGGGGATGCTATGCTGCTATGTT 2012		
QY	481 TCTATTTTGGTGAAGGAAGAAGATCTTTATGATGAAGATACAAATTCGACTATCCATT 540		
DB	2013 CAAATCTTCTTAATGAATCGAAGA---AGAAATGAAGATGTCATCAAGTTCAGCTTG 2069		
QY	541 ATTGACGACCAACAGTAGGTAATCAAGTTTAGTAAATGCTATTTAGTGTCAAGATCGC 600		
DB	2070 ATTGTCGTCCAAAGCTTGGTAAATCATCTTATCAATGCCATCTTCTGTAAGACCGA 2129		
QY	601 GTTATCGTTCTATTTGTCAGGACCAACGAGAGACCTTATTTGATACAGATGTA---TAGT 657		
DB	2130 GTGATTGCAAGCCCAATCGCGGGAACCACTCGTGTGCTCATTTGATACACACTTCGTTGAT 2189		
QY	658 TATGATGCAAGATGATTATTTTAAATCGATACCTGCTGTATGCTGTAATGAAGGAAAGTA 717		
DB	2190 TCTGAAGACCAAGGATTTGTCATGATTTGATACCGCTGGATGCTGAATCAGGAAAT 2249		
QY	718 TATGAATCACTGAGAAATATTCATGATTTAAGAGCTTTAAACGATTTGAAGCTTCAAT 777		
DB	2250 TACGAAACAGTGAAATATTTTATGATTTATCGTGTGCTGCTGCTGCTGCTGCTGCTGCA 837		
QY	778 GTTGTGTTAGTGTATTGATGTCAGAACAGGATTTATGAACAGATTAACGCTGTGCA 837		
DB	2310 ATCGTTTGTGATTAATGCTGGAAGAGGAATTCGCGCAATATCATATGCTGCTATCGCT 2369		
QY	838 GGATATGCACATGAACAGGTAAAGCAGTCTGATTTCTGTAATTAATTAATGGATCTGTTG 897		
DB	2370 GGATTTGCTCATGAGCAGGTAAAGGATTTCTGATTTCTGTAATTAATTAATGGATCTCTT 2429		
QY	898 GAAGAAGTAGTAAACAGATGAAGAAATTTGAAGATGAAGTACGTAAGAAATTTCCAAATTT 957		
DB	2430 GAAGAAGCAATGATACATGAAGAAATTTGAATTAAGAAATTTGTAAGAAATTTCAAAATTC 2489		
QY	958 TTAGATTATGCAAAATTTGCTTTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1017		
DB	2490 CTTGATTATGCAACCAATCGTTTACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2549		
QY	1018 TTCCCTTACATTAATGAAGCAAGTGAAGAAACATGAAGAAACGCTGCTGCTGCTGCTGCTGCTGCT 1077		
DB	2550 CCAGATATGATTAAGAAATTTCCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2609		
QY	1078 AATGAAGCTGTTACTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1130		
DB	2610 AATGATGATGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2662		
RESULT 2			
LOCUS	AI390536		
DEFINITION	552 bp mRNA linear EST 15-MAR-2000		
	mus83b02.y1 Stratagene mouse melanoma (8937312) Mus musculus cDNA		
	clone IMAGE:652107.5, similar to SH:YPHC_BACSU P50743 HYPOTHETICAL		
	48.8 KD GTP-BINDING PROTEIN IN CNK-GPSA INTERGENIC REGION. ;, mRNA		
	sequence.		
ACCESSION	AI390536		
VERSION	AI390536.1 GI:4216543		
KEYWORDS	EST.		
SOURCE	Mus musculus (house mouse)		
ORGANISM	Mus musculus		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
	1 (bases 1 to 552)		
	Marta, M., Hillier, L., Kucaba, T., Martin, J., Beck, C., Wylie, T.,		
	Underwood, R., Steptoe, M., Theising, B., Allen, M., Bowers, J., Person		
	, B., Swaller, T., Gibbons, M., Pape, D., Harvey, N., Schurk, R., Ritter		
	, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann, R.,		
	Waterston, R. and Wilson, R.		
TITLE	The WashU-NCI Mouse EST Project 1999		
JOURNAL	Unpublished		
COMMENT	Contact: Marra M/WashU-NCI Mouse EST Project 1999		
	Washington University School of Medicine		
	4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA		
	Tel: 314 286 1800		
	Fax: 314 286 1810		
	Email: mouseest@wustl.edu		
	This clone is available royalty-free through LLNL ; contact the		
	IMAGE Consortium (infoimage.llnl.gov) for further information.		
	This read is a RESEQUENCE of a previously sequenced mouse clone		
	This read has been verified (found to hit its original self in the		
	correct orientation)		
	Possible reversed clone: similarity on wrong strand		
	MGI:397955		
	Seq primer: -40RP from Gibco		
	High quality sequence stop: 420		
FEATURES			
source	Location/Qualifiers		
	1..552		
	/organism="Mus musculus"		
	/mol_type="mRNA"		
	/db_xref="taxon:10090"		
	/clone="IMAGE:652107"		
	/tissue_type="melanoma"		
	/dev_stage="M2 cells"		
	/lab_host="SOLR (kanamycin resistant)"		
	/clone_lib="Stratagene mouse melanoma (#937312)"		
	/note="Organ: skin; Vector: pBluescript SK-; Site_1: EcoRI		
	; Site_2: XhoI; Cloned unidirectionally. Primer: Oligo		
	dt. From M2 cells, a highly metastatic derivative of the		
	K-1735 (mouse) melanoma. Average insert size: 1.0 kb;		

Uni-ZAP XR Vector: -5' adaptor sequence: 5' CAATCGGCACGAG
3' -3' adaptor sequence: 5' CTCGAGTTTCTTTTCTTTTCTTTT 3'
BASE COUNT 220 a 63 c 94 g 174 t 1 others

Query Match 11.0% Score 143.6; DB 9; Length 552;
Best Local Similarity 56.1%; Pred. No. 4.6e-19;
Matches 316; Conservative 0; Mismatches 235; Indels 12; Gaps 2;

QY 136 GGTGAGTGGTTAAACACATGATTTCAATATTATTATGATACAGGTGGTATTGAATTCGGTCAT 195
DB 1 GTTGATGAGCAGGAAGGAATTTATCTCATAGACAGCGGGAATTTGAATGAAGAAC 60
QY 196 GCACATTTCCAAACACAAATTAGAGCGCAGCGAAGAAATGCCATAGATGAAGCGGATTT 255
DB 61 AAATCATTTCAAGAACAAATTAGAATTTCAAACTCAATAGCAATAGCAAGAACAAATTTA 120
QY 256 ATTATTTTATGTTTAACTGCGTGAAGGATTCACACAAAGCGATGAATGGTGGCTCAA 315
DB 121 ATTATTTTATGATAGTGTAGAGTTGAATAGATTCAGATGATCATTTTGTATTATAGT 180
QY 316 ATTTTATACAAATCTAAACACCGTGTATTAGCGGTGTAAACAAAGTAGATTAATAGGAA 375
DB 181 TTGCTTCGAAATCAAGCAAAAGTTTAACTCGCAGCTAATAAATAGAGGAAGAACAA 240
QY 376 ATGCGTACAGACGTGTATGTTTCTATTATTAGGATTTGGTGAACCGGTATCCAAATATCA 435
DB 241 TTTTCTT-----GATACCTCAATTTTACTAGTGTGTTTGAACATATTTTCCAAATTTCA 294
QY 436 GGGTCACATGTTTAGTCTGTGGTACTGTTAGATGCGAGTTGTTCTCATTTTGGTGA 495
DB 295 GCTATTCAGAGAGAGTGGTGATTTATTAGTAGAACAATAGAAATTTAGATTTT 354
QY 496 GAGGAAGAAGATCTTATGTAAGATACAAATTCGACTATCCATATTCGAGCAGCAAC 555
DB 355 ACAAGA-----GAAGGAAATGAAGCATTTTAGTGGCAATTTAGGAAGAACCAAC 408
QY 556 GTAGTAATCAAGTTAGTAATGCTATTTTAGTGAAGATCGCGTATCTGCTCTTAAT 615
DB 409 GCTGGAATCTACTTTTAAATTAATAGCAATGAATCGCTCCATCTTCTTCTCA 468
QY 616 GTTCGAGGCAACAGAGACGCTATTGATACAGATATGATTATGATGAGCAAGATAT 675
DB 469 ATAGCAAGAACACAGCTGATTCAGTTCTTAGTTTATTAAATTTGCAAAATAGATNTT 528
QY 676 GTTTTAAATCGATCTGCTGGTAT 698
DB 529 GAAATTATAGATCTGCTGGAAT 551

RESULT 3
BJ573090/c 721 bp mRNA linear EST 18-DEC-2002
LOCUS Ipomoea nil mixture of flower and flower bud Ipomoea nil
DEFINITION cDNA clone jn22ml3 3', mRNA sequence.

ACCESSION BJ573090
VERSION BJ573090.1 GI:27254918
KEYWORDS EST.

SOURCE Ipomoea nil (Japanese morning glory)
ORGANISM Ipomoea nil
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons: core eudicots;
asterids; lamids; Solanales; Convolvulaceae; Ipomoea.

REFERENCE 1 (bases 1 to 721)
AUTHORS Hoshino,A., Seki,M., Shin-i,T., Carninci,P., Kamiya,A., Shiraki,T.,
Nitasaka,E., Shinozaki,K., Hayashizaki,Y., Kohara,Y. and Iida,S.
TITLES ESTs of Japanese morning glory
JOURNAL Unpublished

COMMENT Contact: Tadasu Shin-i
Center For Genetic Resource Information
National Institute Of Genetics
1111 Yata, Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6856

Fax: 81-559-81-6855
Email: tshin@genes.nig.ac.jp.
Location/Qualifiers
1. .721
/organism="Ipomoea nil"
/mol_type="mRNA"
/cultivar="Tokyo-kokei standard"
/db_xref="taxon:35883"
/clone="jm22ml3"
/tissue_type="mixture of flower and flower bud"
/clone_lib="Ipomoea nil mixture of flower and flower bud"

BASE COUNT 186 a 184 c 136 g 215 t
ORIGIN

Query Match 9.8% Score 128.4; DB 12; Length 721;
Best Local Similarity 52.8%; Pred. No. 6.2e-16;
Matches 378; Conservative 0; Mismatches 321; Indels 17; Gaps 4;

QY 587 TAGGTGAGATCGCTTATCGTTTCTAATGTTGCAGGCAACAGAGAGCGCTATTGATA 646
DB 719 TTGGGAGAGACAGACAATTTGTTAGTCCCGTCAGTGGAACTACTCTGATGCAATTTGATA 660
QY 647 CAGAGTATA---GTTATGATGACAAAGATTATGTTTAAATCGATACCTGCTGGTATCGGTA 703
DB 659 CAGAGTTTACTGGATCAGATGACAGAAATATCGTCTCATTTGATGATCTGCAGGAATAAGGA 600
QY 704 AAAAGGAAAAGATATATCAATC-----AACTGAGAAATATTTCAGTATTAAAGAGCTT 754
DB 599 AAAGGCAACTCTGCTTCATCTGTAGCATAACAGAGCTTTATCTGTAATCAAGCAT 540
QY 755 TAAAGCGATTCAGCTTCAATGTTGTTTAGTGGTATTGATGCAAGAACAGGCAATTA 814
DB 539 TCGTGCAATAGCTGTCAGATGTTGGCTCTTGTAAATGAGGCTATGCTCTTGCATCA 480
QY 815 TTGAACAAAGATAAAGCTGTTCAGGATATGCACATGAACAGGTAAGAGAGTCTGATTCG 874
DB 479 CTGAGCAGGATTCGAAGATTCGAGAAAGATAGAGAGGAGGAGGATTTGCTCTGCTTCG 420
QY 875 TCGTAATAAATGGGATCTGTGGAAAGATAGATAA--AACGATGAAGAAATTTCAAGA 932
DB 419 TTGTAATAAGTGGGACACAATTCACACAAAACCAATCACTGCAACTTACTATGAGCA 360
QY 933 TGAAGTACGTAAGAATTTCCAAATTTTAGATATTGCAAAATTTGCTTTTGTGCTGCTAA 992
DB 359 AGATGTTAGGGGGAAGCTCCGTTGTTGGTTGGGCACCTATTGTTGATTCAACTGCAAT 300
QY 993 AGAACGCACAAGATTACGTATATTCCTTTACATTAATTAAGCAAGTGAAACCACTAA 1052
DB 299 AGCAGGGCATACGCTCGACAAAATTTATTTGTTACTGCTGCTGTTGAAAAAGAGATC 240
QY 1053 AAAACGTGTTCAAGTTCAACTTTAAATGAAGTCTTACTGATGCAATTTCCATGAACCC 1112
DB 239 AGAAGATTAAACCGGCAATTCGAATCAAGTGGTAGCGGAGGAGCTGCGGTTCAAATC 180
QY 1113 TACACCAACAGACAAAGGT---AGACGTTTGAATGCTCTTTTATGCAACACAAAGTTCTGTAT 1169
DB 179 ACCTCCAGGACTAGAGGTGGGAGAGGGGTCGTATATTATTGCACTCAGGCTGCTAT 120
QY 1170 AGAACCCAGCAGATTTGTTGTTATTTGTTAATGATGATAGAAATTAATGCAATTTTCTTATAA 1229
DB 119 TAGACCACTACTTTTGTCTTTGTCATGATGCAAAACTTTTCTCCGAGACATACCG 60
QY 1230 ACCTATTTTAGAATCAAAATCCGTCGCGCTTTTGGCTTTTGAAGGTACACCAATTC 1285
DB 59 ACGGTACATGGAAAAAGCAACTGAGATCAAGTGCAGGCTTCTCGGCGCACACCTATTC 4

RESULT 4
AV110834 1233 bp mRNA linear HTC 17-OCT-2002
LOCUS Zea mays CL915_1 mRNA sequence.
DEFINITION AV110834
ACCESSION AV110834.1 GI:21215424

LOCUS BQ514303 579 bp mRNA linear EST 07-MAR-2003
 DEFINITION EST621718 Generation of a set of potato cDNA clones for microarray analyses mixed potato tissues Solanum tuberosum cDNA clone STW112 5' end, mRNA sequence.

ACCESSION BQ514303
 VERSION BQ514303.1 GI:21373172
 KEYWORDS EST.
 SOURCE Solanum tuberosum (potato)
 ORGANISM Solanum tuberosum

REFERENCE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Asterids; lamids; Solanales; Solanaceae; Solanum.
 1 (bases 1 to 579)
 Buell, C.R., Hart, A., Baker, B., Tanksley, S., Fry, W., Smart, C., Restrepo, S., Griffiths, H., van der Hoeven, R., Tsai, J., and Karamycheva, S.A.

AUTHORS Generation of a set of potato cDNA clones for microarray analyses
 Unpublished
 Other ESTs: EST621719
 Contact: Robin Buell
 The Institute for Genomic Research
 9712 Medical Center Dr, Rockville, MD 20850, USA
 Email: potato-array@tigr.org
 This clone can be obtained from the University of Arizona Genomics Institute. Orders can be made through URL:
 http://genome.arizona.edu/orders/
 Seq primer: T3

TITLE Location/Qualifiers
 COMMENT 1..579
 /organism="Solanum tuberosum"
 /mol_type="mRNA"
 /cultivar="Kennebec or Binjite"
 /db_xref="taxon:4113"
 /clone="STW112"
 /tissue_type="mixed tissues"
 /lab_host="SOLR"
 /clone_lib="Generation of a set of potato cDNA clones for microarray analyses mixed potato tissues"
 note="vector: Bluescript SK(-); Site.1: EcoRI; Site.2: XhoI; supplier: Combination of untreated and Phytoththora infestans-treated libraries of stolons, leaves, leaflets, axillary buds of stem explants, petioles, germinating eyes tubers, or roots."

BASE COUNT 179 a 95 c 153 g 152 t
 ORIGIN

Query Match
 Best Local Similarity 55.7%; Pred. No. 4.2e-10;
 Matches 238; Conservative 0; Mismatches 177; Indels 12; Gaps 2;

QY 531 ACTATCCATTATTGGACGACCAACCTAGCTAATCAAGTTTAGTAAATGCTATTTAGG 590
 DB 21 AATCGCAATTGTTGGCGCGTCCAAATGTTGTAAGTAGTATTTTGAATGCTTTAGTCGG 80

QY 591 TGAAGATCGGTTATCTGTTCTTAATGTTGACGGACCAACGAGACGCTATTGATACAGA 650
 DB 81 AGAGGACAGACAATTTAGTCCCGTTAGTGGAACTACTCGTAGCTATAGATACAGA 140

QY 651 GT---ATAGTATGATGACAGCAAGATTATCTTTATTCGATCTGCTGATGCGTAAAAA 707
 DB 141 ATTCATGTCAGACAGATGGCGAAGTTTCGGGTTATAGATACCTGCCGAATTAAGAAAA 200

QY 708 AGGAAAAGTATATGAATC-----AACTGAGAAATATTACGATTATTAAGAGCTTTAAA 758
 DB 201 GACTGCGACTGCTCATCTAGGTAGCATACCAAGCGCTTTATCAGTAATCAAGCATTCG 260

QY 759 AGCGATTGAACGTTCAATGTTGTTTGTAGTGGTTATTGATGACAGAACCAAGCATTTGA 818
 DB 261 TGCATTCGTCGTCAGATGTTGCTCTTCTTATTGAGGCTTAGGCTTGCTGCATCATGA 320

QY 819 ACAAGATAAAGCTGTTGACGAGATATCCATGCAACAGGTAAGACAGTCGATGATTCGT 878
 DB 321 ACAGGATTGCAAAATTGCAAGAGGATAGAGAGAGGAGGAGGCTTCTCTCATCGTTGT 380

QY 879 AAATAAATGGGATCTGTGGAAAAAGATAGTAAACCATGATGAACAAATTTGAAGATGAAGT 938
 DB 381 GAACAAGTGGGATACGATCCCAACACAGAACCAAGAACTACTGTATTTTATGAGGAAGA 440

QY 939 ACGTAAA 945
 DB 441 TGTAGA 447

RESULT 8
 BQ281519
 LOCUS BQ281519
 DEFINITION WHE3022_D01_G02S wheat unstressed seedling shoot normalized cDNA library Triticum aestivum cDNA clone WHE3022_D01_G02, mRNA sequence.

ACCESSION BQ281519
 VERSION BQ281519.1 GI:20549073
 KEYWORDS EST.
 SOURCE Triticum aestivum (bread wheat)
 ORGANISM Triticum aestivum

REFERENCE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Poideae
 1 (bases 1 to 532)
 Anderson, O.D., Chao, S., Choi, D.W., Close, T.J., Fenton, R.D., Han, P.S., Hsia, C.C., Kang, Y., Lazo, G.R., Miller, R., Nguyen, H.T., Rausch, C.J., Seaton, C.B., Tong, J.C., and Zhang, D.

AUTHORS The structure and function of the expressed portion of the wheat genomes - Normalized shoot cDNA library
 Unpublished
 Contact: Olin Anderson
 US Department of Agriculture, Agriculture Research Service, Pacific West Area, Western Regional Research Center
 800 Buchanan Street, Albany, CA 94710, USA
 Tel: 5105595773
 Fax: 5105595818
 Email: oanderson@pw.usda.gov
 Sequences have been trimmed to remove vector sequence and low quality sequence with phred score less than 20
 Seq primer: SK primer.

TITLE Location/Qualifiers
 JOURNAL 1..532
 COMMENT /organism="Triticum aestivum"
 /mol_type="mRNA"
 /cultivar="Chinese Spring"
 /db_xref="taxon:4565"
 /clone="WHE3022_D01_G02"
 /tissue_type="Etiolated shoot"
 /dev_stage="Five day old seedling"
 /lab_host="E. coli DH10B"
 /clone_lib="Wheat unstressed seedling shoot normalized cDNA library"

note="vector: Lambda Uni-ZAP XR, excised phagemid pBluescript SK; Site.1: EcoRI; Site.2: XhoI; Seeds were surface-sterilized, germinated and grown aseptically in the dark at room temperature on filter paper with water, nystatin and cefotaxime in covered crystallization dishes. Roots were harvested. The tissue, total RNA, and poly(A) RNA were prepared, a cDNA library was made in the pT7 clone lab (Choi, Close, Fenton) at the University of California, Riverside. The cDNA clones were in vivo excised to give phagescript phagemids before normalization was carried out. The mass excision of phagemid library and normalization were done in HT Nguyen lab by D. Zhang at Texas Tech University. Normalization protocol used was that of Soares'. Plasmid DNA preparations and DNA sequencing were performed in the OD Anderson lab (all other authors)."

BASE COUNT 162 a 112 c 110 g 148 t
 ORIGIN

Query Match
 Score 99.2; DB 13; Length 532;

```

Best Local Similarity 53.3%; Pred. NO. 5.5e-10; Mismatches 216; Indels 6; Gaps 2;
Matches 256; Conservative 0;

QY 753 TTTAAAGCGATTGAACGTTCAAAATGTTTGTAGTGGTATTATGTCAGAACAGGCAT 812
Db 48 TTTTCTAGCAATTCGCGTCTGATGTTGCTTGTATCTAGGCAATGGCGTGTGT 107
QY 813 TATTGAACAAGATAAAGCTGTTGTCAGGATATGACATGAACAAGTAAGCAAGTCGTCAT 872
Db 108 TACCGACGAAGATTATATAATTCACAGAGGATTGCAAGAAGGAAAGCTTGTGTCAT 167
QY 873 TGTGTAATTAATGCGTACTGTGGAAGAAGATA---GTAAACGATGACAAATTTGA 929
Db 168 TGTGTGAACAATGGGATGATATCCCTAACACAGACCATGAGATACACACATATGA 227
QY 930 AGATGAATACGTAAAGAAATTCCAATTTTATGATATGACAAATTTGTTGTCTGC 989
Db 228 ACAAGATGTAATGAAGAAGCTTCCCATCTGATTGGCACCCTATTGTCTACTGTTCTGC 287
QY 990 TAAAGAAGCGACGAAGATTAGCTACATTATTCCTTACATTAATGAAGCAAGTGAACCA 1049
Db 288 GACTAATGGCAGCAGTGTCAAAAGATATTTCTGCTGCTGCTGTTGTTGAGATGAAG 347
QY 1050 TAAAAACGTTTCAAGTTCAACTTTAAATGAAGTCGTACTGATGCAATTTCCATGAA 1109
Db 348 GTCTATAGACTCGGCACCTCCATTTCTTAATCAAGTTATTAGAGAGCTATAGCATTCAA 407
QY 1110 CCCTACACCAACAGACAAAGT---AGACGTTTGAATGCTTTTATGCAACACAGTTGC 1166
Db 408 ACCACCAACGAAGAGGTCGCAAAAGAGGTCGTCTCTATTATACACACAGGCTGC 467
QY 1167 TATGACACCAACGATTTTCTGATTTTATGATGATGATGATGATGATGATGATGAT 1226
Db 468 CATTGGTCCCAACATTCGTTTATTTTGTAAATGATGGAAGCTCTTCCCTGATACATA 527

RESULT 9
AI465410/c
LOCUS
DEFINITION
mu83b02.x1 Stratagene mouse melanoma (#937312) Mus musculus cDNA
clone IMAGE:652107 3', similar to SW:PHC-BACSU P50743 HYPOHETICAL
48.8 KD GTP-BINDING PROTEIN IN CMK-GPSA INTERGENIC REGION. ; mRNA
sequence.
ACCESSION
AI465410
VERSION
AI465410.1 GI:4319440
SOURCE
Mus musculus (house mouse)
ORGANISM
Mus musculus
REFERENCE
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 395)
AUTHORS
Marras M., Hillier L., Kucaba T., Martin J., Beck C., Wylie T.,
Underwood K., Stepien M., Theising B., Allen M., Bowers Y., Person
B., Swaller T., Gibbons M., Pape D., Harvey N., Schurk R., Ritter
E., Kohn S., Shin T., Jackson Y., Cardenas M., McCann R.,
Waterston R. and Wilson R.
The WashU-NCI Mouse EST Project 1999
Unpublished
TITLE
Contact: Marra M/WashU-NCI Mouse EST project 1999
JOURNAL
Washington University School of Medicine
COMMENT
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@wustl.edu
This clone is available royalty-free through LLNL ; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
This clone was previously sequenced on the 5' end only, this new
data is from the 3' end
Possible reversed clone: similarity on wrong strand
High quality sequence stop: 276.
Location/Qualifiers
1..395
/organism="Mus musculus"

/mot_type="mRNA"
/db_xref="taxon:10090"
/clone="IMAGE:652107"
/tissue_type="melanoma"
/dev_stage="M2 cells"
/lab_host="SOLR (kanamycin resistant)"
/clone_lib="Stratagene mouse melanoma (#937312)"
/Note="Organ: skin; Vector: pBluescript SK-; Site: 1: EcoRI
Site: 2: XhoI; Cloned unidirectionally. Primer: Oligo
dT. From M2 cells, a highly metastatic derivative of the
K-1735 (mouse) melanoma. Average insert size: 1.0 kb;
Uni-ZAP XR vector; -5' adaptor sequence: 5' GAATTCGCCACGAG
3' -3' adaptor sequence: 5' CTCGAGTTTTTTTTTTTTTTT 3'."

BASE COUNT 140 a 41 c 47 g 167 t
ORIGIN
Query Match 6.9%; Score 89.4; DB 9; Length 395;
Best Local Similarity 55.3%; Pred. NO. 5.3e-08;
Matches 194; Conservative 0; Mismatches 156; Indels 1; Gaps 1;
QY 955 TTTTATGATTTATGACAAATGCTTTGCTGCTTAAAGACGACAGATGATGATCA 1014
Db 364 TTTTATGCTGAGCCCAATTTGTTTATTTCAGC-ACACAGGAAGTAGAATTTCTAAA 306
QY 1015 TTTATCCCTTACATTAATGAAGCAAGTGAAGCAATTAAGCAATTAAGCAATTAAG 1074
Db 305 TTAAGAGAACCAATTTTAGAAGTCAAAATTAATTAAGTAAGTAAGTAAGTAAGT 246
QY 1075 TTAATGAAGTCGTTTACTGATGCAATTTCCATGACCCCTACACACACAGAGTGA 1134
Db 245 CTTAATCAATTCATTTAGATATTCATTAATGTTTACACCTCCCAAGTCTTCTG 186
QY 1135 CTTTGAATGCTTTTATGCAACACAGTTGCTATAGAACCCGACATTTGTTGATTT 1194
Db 185 AAATTAATTAATTAATTTGCTCAACAAATTAATTAATTAATTAATTAATTAATTT 126
QY 1195 GTTAATGATGATTAATTAATGCTTTTCTTATAAACGCTATTATTAGAGATCAAT 1254
Db 125 GTTAATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 66
QY 1255 GCGCTTTGTTGTTGAGGTACACCAATTCATATTATAGCTGGAAGAGA 1305
Db 65 AACTATTTGTTTTCAGGTGTCCTCAATTAATTAATTAATTAATTAATTAATTAATTA 15

RESULT 10
CD377507
LOCUS
DEFINITION
PTM02676 Phaeodactylum tricornutum Uni-zap XR Phaeodactylum
tricornutum cDNA 5', mRNA sequence.
ACCESSION
CD377507.1 GI:31253121
VERSION
CD377507
KEYWORDS
Phaeodactylum tricornutum
SOURCE
Phaeodactylum tricornutum
ORGANISM
Eukaryota; stramenopiles; Bacillariophyta; Bacillariophyceae;
Bacillariophycidae; Naviculales; Phaeodactylaceae; Phaeodactylum.
1 (bases 1 to 745)
AUTHORS
Scala S., Carls N., Falciatore A., Chiusano M.L. and Bowler C.
TITLE
Genome properties of the diatom Phaeodactylum tricornutum
JOURNAL
Plant Physiol. 129 (3), 993-1002 (2002)
MEDLINE
2211123
PUBMED
12114555
COMMENT
Contact: Bowler C
Laboratory of Molecular Plant Biology
Stazione Zoologica 'Anton Dohrn'
Villa Comunale, I-80121, Napoli, Italy
Tel: 39 081 583 3268/3211
Fax: 39 081 764 1355
Email: chris@lephs.zsn.it
Diatom EST Database(http://avesthagen.szbowlwer.com)
Seq primer: T3 backward
POLYA-No.

```



```

/organism="Drosophila melanogaster"
/mol_type="genomic DNA"
/db_xref="taxon:7227"
/clone="BACR08K10"
/clone_lib="RPCI-98"
/note="end : TET3"
BASE COUNT      201 a      64 c      131 g      202 t      503 others
ORIGIN

Query Match      6.4%; Score 83.2; DB 29; Length 1101;
Best Local Similarity 19.6%; Pred. No. 1.1e-06;
Matches 130; Conservative 271; Mismatches 258; Indels 3; Gaps 1;

QY 441 ACATGGTTAGCTTGGTGGACATTTAGTATGATGCGATGTTCTTCATTTTGGTGAAGAGGA 500
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1099 ARRGGDDTWDRTKDDWONTKWTWTKDRDORRMAGDADRWAMDDGAGTMTATMMW 1040

QY 501 AGAAGATCCTTATGATGAAGATCAATTCGACTATCCATTTATTCGACCAACACGTAGG 560
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1039 WMMWATWDTWKKMMWATAAKTDTATWMTAWRADWAGRGAGKRDRAATDADGAG 980

QY 561 TAATCAAGTTTAAATGCTATTTTAGTGAAGATCGCTTATCTCTTCTATGTTGC 620
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 979 RRGGRKKDKKDDDDKGGKKKAAKAAKAWATKMDDDMDKMDKWDGAKDRAD 920

QY 621 AGGCACACGAGAGCGCTATTGATACAGATATAGTTATGATGACAGAGATATGTTTT 680
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 919 DDDGAGDKDDGKDDDDTGTGTDKDDKDDKDDKAKGTGWDATWAAATDMMWGW 860

QY 681 AATCGATCTGCTGGTATGCTGTAAGAAAGAAAGATATATGAATCAACTGAGAATATTC 740
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 859 ADADWTTWDAADDWADDDWADWANKWDDANWARTADRRDWDGAGKGGARKRRDR 800

QY 741 AGTATTAAAGCGCTTTAAACCGGATTCGAACGTTCAAAATGTTGTTTAGTGGTTATGATGC 800
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 799 KRADDKDAADDKDAATWTTTWTTR---DTDDKWKKTDTWTRWAADRTWDRDDDD 743

QY 801 AGAACAGGCTATTATGAACAGATAAACGCTGTCGAGCATATGCGATGACATGAACAGGTAA 860
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 742 DRAGTAGKRRRTWKRWRKRRDRDADAADDTARDRRRRRGDDGADAGKCKTKGRKRR 683

QY 861 AGCAGCTGCTATGCTGTAATAAATGGCATCTGCGAANAAGATAGTAAACGATGAA 920
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 692 RDRATWDTDAWADAANWTTTDTDDMDKDRRRKGAARRRTTARAADWMTWRAWD 623

QY 921 GAAATTTGAAGATGAAGTACGTAAAGAAATTCCAATTTTATGATATGACAAATGCTTT 980
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 622 WAKDWKTRADRWDAADTWTDAKADRWAKARARRRDRARAARADRRWTTKGKTT 563

QY 981 TGTCTGCTGCTAAAGACGACAGATTCGATATTTCCCTTACATTAATGAAGCAAG 1040
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 562 TATWTTWAARAANWAWANATTTATWTTTWTWTTTWTWTTTWTWTTTAAWNAWTA 503

QY 1041 TGAACACATAAAGCGTGTTCAAAGTTCAACTTTAAATGAAGTCGTTACTGATGCAAT 1100
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 502 TAAATTAATAAANAANAANAATTTTTTTTTTTTAAWNAWMTAATWTTTWTWTTTAAAT 443

QY 1101 TT 1102
|
Db 442 TT 441

RESULT 13
BH234795
LOCUS
DEFINITION
MEAA_B03.y E Spiroplasma kunkelii genomic clone MEAA_B03.y, genomic survey sequence.
ACCESSION
BH234795
VERSION
BH234795.1 GI:18030263
KEYWORDS
GSS.
SOURCE
Spiroplasma kunkelii
ORGANISM
Bacteria; Firmicutes; Mollicutes; Entomoplasmatales;

Spiroplasmataceae: Spiroplasma.
1 (bases 1 to 167)
Hogenhout,S.A.
Genomic sequences from Spiroplasma kunkelii strain M2
Unpublished
Contact: Hogenhout SA
Department of Entomology
The Ohio State University-OARDC
120 Thorne Hall, 1680 Madison Avenue, Wooster, OH 44691, USA
Tel: 330 263 3730
Fax: 330 263 3686
Email: hogenhout.1@osu.edu
Plate: AA row: B column: 03
Class: EcoRI fragments.
Location/Qualifiers
1..167
/organism="Spiroplasma kunkelii"
/mol_type="genomic DNA"
/strain="M2"
/db_xref="taxon:47834"
/clone="MEAA_B03.y"
/clone_lib="E"
BASE COUNT      58 a      26 c      26 g      57 t
ORIGIN

Query Match      6.2%; Score 81.4; DB 28; Length 167;
Best Local Similarity 68.7%; Pred. No. 2.1e-06;
Matches 112; Conservative 0; Mismatches 51; Indels 0; Gaps 0;

QY 533 TATCATTTATGGAGCACCACCAAGTAGTAATCAAGTTTAGTAATGCTATTTAGGTG 592
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 5 TTTCTTTAATCGGTAAACCAATGTTGCAAAATCTCTTTTACAAATGCTATTTAGGTG 64

QY 593 AGATCGCGTTATCGTTTCTTAATGTTCCAGGACAGAGACGCTATTTGATACAGAGT 652
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 65 AAGAACGAGTTTGTATCATCCCAATTCAGGAACAACACTAGATTCAATTGATACATCAT 124

QY 653 ATAGTTATGATGACGACAGATATGTTTAAATCGATCTGCTGG 695
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 125 TTAAACGAATAAATCAACTTTTATATCTGTGATTGATCTGCTGG 167

RESULT 14
BQ55888
LOCUS
DEFINITION
QGB28A05.yg.ab1 OG-ABCDI lettuce salinas Lactuca sativa cDNA clone
ACCESSION
BQ55888
VERSION
BQ55888.1 GI:22241353
KEYWORDS
EST.
SOURCE
Lactuca sativa
ORGANISM
Lactuca sativa
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; campanulids; Asterales; Asteraceae; Cichorioideae;
Cichoraceae; Lactuca.
1 (bases 1 to 653)
Knapp,S., Matvienko,M., Lavelle,D., Chevallier,P., Ziegler,J., Ellison
Lin,H., van Damme,M., Slabaugh,M.S., Livingston,K., Zhou,Y., Lai,Z.,
Church,S., Jackson,L. and Bradford,K.
Lettuce and Sunflower ESTs from the Compositae Genome Project
http://compgenomics.ucdavis.edu/
Unpublished
Contact: Alexander Kozik [R.W.Michelmore]
Department of Vegetable Crops, R.W.Michelmore Lab
University of California at Davis (UCD)
Asmundson Hall, UCD, Davis, CA 95616, USA
Tel: 1-(530)-742-1742
Fax: 1-(530)-752-9659
Email: akozik@ucdavis.org [michelmoreveg@mail.ucdavis.edu]
singleton; see http://cgdb.ucdavis.edu/ for details.
Plate: QGB28 row: A column: 05.
```


GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 12, 2003, 17:17:09 ; Search time 5327.98 Seconds
(without alignments)
10020.134 Million cell updates/sec

Title: us-09-815-242-4228

Perfect score: 1305

Sequence: 1 agcactaacctatagtagc.....atattatagctgaaagaga 1305

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2888711 seqs, 20454813386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0
Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl:*

- 1: gb_ba.*
- 2: gb_htg.*
- 3: gb_in.*
- 4: gb_om.*
- 5: gb_ov.*
- 6: gb_pat.*
- 7: gb_ph.*
- 8: gb_pl.*
- 9: gb_pr.*
- 10: gb_ro.*
- 11: gb_sts.*
- 12: gb_sv.*
- 13: gb_un.*
- 14: gb_vl.*
- 15: em_ba.*
- 16: em_fun.*
- 17: em_hum.*
- 18: em_in.*
- 19: em_mu.*
- 20: em_om.*
- 21: em_or.*
- 22: em_ov.*
- 23: em_pat.*
- 24: em_ph.*
- 25: em_pl.*
- 26: em_ro.*
- 27: em_sts.*
- 28: em_un.*
- 29: em_vl.*
- 30: em_htg_hum.*
- 31: em_htg_inv.*
- 32: em_htg_other.*
- 33: em_htg_mus.*
- 34: em_htg_pln.*
- 35: em_htg_fod.*
- 36: em_htg_mam.*
- 37: em_htg_vrt.*
- 38: em_sy.*
- 39: em_htgo_hum.*
- 40: em_htgo_mus.*
- 41: em_htgo_other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
C 1	1292.2	99.0	301550	1	AP003134	AP003134 Staphyloc
C 2	1292.2	99.0	346900	1	AP003362	AP003362 Staphyloc
C 3	1290.6	98.9	1308	6	AX622668	AX622668 Sequence
C 4	1290.6	98.9	333750	1	AP004827	AP004827 Staphyloc
C 5	919.4	70.5	3269	1	AF270032	AF270032 Staphyloc
C 6	919.4	70.5	3269	6	AX145350	AX145350 Sequence
C 7	919.4	70.5	300892	1	AE016747	AE016747 Staphyloc
C 8	733.8	56.2	304680	1	AE017002	AE017002 Bacillus
C 9	727.4	55.7	290117	1	AE017028	AE017028 Bacillus
C 10	683.2	52.4	347050	1	AL591981	AL591981 Listeria
C 11	683.2	52.4	349980	6	AX641670	AX641670 Sequence
C 12	676.8	51.9	313450	1	AL596170	AL596170 Listeria
C 13	676.8	51.9	319630	6	AX413016	AX413016 Sequence
C 14	676.8	51.9	349980	6	AX417046	AX417046 Sequence
C 15	667.6	51.2	960	6	AX144037	AX144037 Sequence
C 16	656.8	50.3	300050	1	AP004599	AP004599 Oceanobac
C 17	647.4	49.6	24887	1	BACSERA	L47648 Bacillus su
C 18	647.4	49.6	213680	1	BSUB0012	Z99115 Bacillus su
C 19	634.8	48.6	302173	1	AE016951	AE016951 Enterococ
C 20	633	48.5	300550	1	AP001512	AP001512 Bacillus
C 21	600	46.0	1311	6	AX607165	AX607165 Sequence
C 22	600	46.0	20501	1	AE014265	AE014265 Streptoco
C 23	600	46.0	44145	6	AX602195	AX602195 Sequence
C 24	600	46.0	174050	1	SAG766852	AL766852 Streptoco
C 25	599.4	45.9	1311	6	AX433884	AX433884 Sequence
C 26	582.4	44.6	11071	1	AE006309	AE006309 Lactococc
C 27	552	42.3	52276	1	AE014141	AE014141 Streptoco
C 28	552	42.3	151947	2	SPNEU1902	AL449524 Streptoco
C 29	552	42.3	323825	1	AP005146	AP005146 Streptoco
C 30	550.4	42.2	1308	6	AX570314	AX570314 Sequence
C 31	550.4	42.2	5066	6	AR218847	AR218847 Sequence
C 32	550.4	42.2	5066	6	BD003759	BD003759 Polynucle
C 33	550.4	42.2	10310	1	AE007464	AE007464 Streptoco
C 34	550.4	42.2	349980	6	AX571764	AX571764 Sequence
C 35	550.4	42.2	349980	6	AX571765	AX571765 Sequence
C 36	549.6	42.1	12434	1	AE006498	AE006498 Streptoco
C 37	548.8	42.1	12370	1	AE009978	AE009978 Streptoco
C 38	548.8	42.1	12540	1	AE008523	AE008523 Streptoco
C 39	544.2	41.7	302050	1	AL935257	AL935257 Lactobaci
C 40	531.2	40.7	3737	1	AB016077	AB016077 Streptoco
C 41	531.2	40.7	13860	1	AE015016	AE015016 Streptoco
C 42	486.2	37.3	301278	1	AE015939	AE015939 Clostridi
C 43	476	36.5	296750	1	AP003191	AP003191 Clostridi
C 44	438.4	33.6	10861	1	AE007680	AE007680 Clostridi
C 45	423.2	32.4	10691	1	AE010530	AE010530 Fusobacte

ALIGNMENTS

RESULT 1
AP003134/c
LOCUS AP003134 301550 bp DNA linear BCT 24-APR-2003
DEFINITION Staphylococcus aureus subsp. aureus N315 genomic DNA, complete
genome, section 6/10.
ACCESSION AP003134 BA000018
VERSION AP003134.2 GI:14349226
KEYWORDS
SOURCE
ORGANISM Staphylococcus aureus subsp. aureus N315
Bacteria; Firmicutes; Bacilliales; Staphylococcus.
REFERENCE
AUTHORS Kuroda, M., Ohta, T., Uchiyama, I., Baba, T., Yuzawa, H., Kobayashi, I.,
Cui, L., Oguchi, A., Aoki, K., Nagai, Y., Iino, J., Ito, T., Kanamori, M.,
Matsumaru, H., Maruyama, A., Murakami, H., Hosoyama, A.,

Mizutani-Ui, Y., Takahashi, N.K., Sawano, T., Inoue, R., Kaito, C.,
 Sekimizu, K., Hirakawa, H., Kuhara, S., Goto, S., Yabuzaki, J.,
 Kanehisa, M., Yamashita, A., Oshima, K., Furuya, K., Yoshino, C.,
 Shiba, T., Hattori, M., Ogawara, N., Hayashi, H., and Hiramatsu, K.
 Whole genome sequencing of methicillin-resistant *Staphylococcus*
aureus
 Lancet 357 (9264), 1225-1240 (2001)
 21311952
 11418146
 2 (bases 1 to 301550)
 Director-General, Biotechnology Center, Aoki, K., Oguchi, A.,
 Hosoyama, A., Nagai, Y., Kuroda, M., Hiramatsu, K. and Kikuchi, H.
 Direct Submission
 Submitted (30-JAN-2001) Director-General, Biotechnology Center,
 National Institute of Technology and Evaluation, Biotechnology
 Center, 2-chome 49-10 Nishihara, Shibuya-Ku, Tokyo 151-0066, Japan
 (E-mail: biotechite.go.jp, URL: http://www.bio.nite.go.jp/
 Tel: 81-3-3481-1933, Fax: 81-3-3481-8424)
 On Jun 12, 2001 this sequence version replaced gi:13701258.
 Location/Qualifiers
 1. 301550
 /organism="Staphylococcus aureus subsp. aureus N315"
 /mol_type="genomic DNA"
 /strain="N315"
 /sub_species="aureus N315"
 /db_xref="taxon:158879"
 complement(249. .947)
 /gene="SA1293"
 complement(249. .947)
 /gene="SA1293"
 /note="ORFID:SA1293"
 /codon_start=1
 /transl_table=11
 /product="conserved hypothetical protein"
 /protein_id="BAB42553.1"
 /db_xref="GI:13701259"
 /translation="MSFTNIIYFVILMVLPLHAQHVKSNYEKYSQVRSCTGKGR
 EVAILHANGITDVEVGEGLDTHDPPKRVVLSLPANDRPVSAGTAIAAEVHG
 EAQDHGGHFLFRALVFPVANGSSVTKHIIIGLITAIAGSAGFTALWIGAGLM
 SLAVFSITLVPFPOASSRANKQITALNIVNEKEYHARKVLAAAAMTYVAATAVAV
 AELVRIILIARSSD"
 complement(1001. .1588)
 /gene="SA1294"
 complement(1001. .1588)
 /gene="SA1294"
 /note="ORFID:SA1294"
 /codon_start=1
 /transl_table=11
 /product="conserved hypothetical protein"
 /protein_id="BAB42554.1"
 /db_xref="GI:13701260"
 /translation="MTINTFWQYTLQRSWLMMLILCNILGMYYGYWYGRQLSHTPW
 HFKIFVDPSTAILYLVISILIKQNSIIEALAFVTLFKYGIWAVIMNIIIFIQ
 GDITVGLVMSHISIMAVQAIYFPRPKRSIGISVAMWVFLNDYIDVFLQFPYY
 DFITTHVQIGVLSCCLSVFGLLXIELNKLKCK"
 complement(1578. .2153)
 /gene="SA1295"
 complement(1578. .2153)
 /gene="SA1295"
 /note="ORFID:SA1295"
 /codon_start=1
 /transl_table=11
 /product="conserved hypothetical protein"
 /protein_id="BAB42555.1"
 /db_xref="GI:13701261"
 /translation="MSETLNQIKESFIEVLLFOYRFRKSRIANVWLNKYVNEAKLANI
 HFVDTKLNHETAEVGSASHASAOFTKRNKLKMTNEIFDIYANHNCAFDIOIHFA
 NYSKREORLDLIAVQUTESPSTQYTHLNLNSAIDRHKHALLIDYLLHNIDLSLQNM
 EKQRFYQTLQILNKLKLNKHNQFEDLADD"
 complement(2167. .3411)
 /gene="SA1296"
 complement(2167. .3411)
 /gene="SA1296"

/note="ORFID:SA1296"
 /codon_start=1
 /transl_table=11
 /product="conserved hypothetical protein"
 /protein_id="BAB42556.1"
 /db_xref="GI:13701262"
 /translation="MEDIYKLIIDDINLQKLENLDSRVNEAITTNDNDALFILGETLYN
 FGLPQGLEVFRVLYHKIPDESELLYFIEGLMSENQDEALEYLSYVPSPEKMLLE
 ADLYQOINMHEVAIDKQEALELEPNDFIHFALAEMLYDQYLRATSEYETVLETG
 EYQVGNVLFERNADCKQEALELEPNDFIHFALAEMLYDQYLRATSEYETVLETG
 AIKMTWLLSKDPDYQGYLYLQSLYENENKYPDAIETGKEGLRSQFYKELMTTTC
 LEIGHDANEGVLLKQAEVONAYQEPILLISDLYNEDYEAIIEILLTVDEEDLD
 PTFMHLAFAYGOERDKEAHEFFELAYPTMTNIDFNSDYFYVLEIGONEQAITLL
 NOLLESPSNENHDELSRLQS"
 complement(3418. .4716)
 /gene="arob"
 complement(3418. .4716)
 /gene="arob"
 /note="ORFID:SA1297"
 /codon_start=1
 /transl_table=11
 /product="3-phosphoshikimate 1-carboxyvinyltransferase"
 /protein_id="BAB42557.1"
 /db_xref="GI:13701263"
 /translation="MSEQIIDIISGPKGIEVPGDKSMTHRAIMLASLAEGETSNLYK
 PLLGECRRRTMDIFRLLCVDIKEDDKLVNSPGYKAEKTHOVLYTGNSTVTRLLA
 GLSLGIGIESVLSGDSIGARPHDRVRLKPLADANIEGIEDNTYPLIKRSPVINGIN
 YQMEVASAQVKSAILFASLFSNODTTVIKELDSRNHTETMFRHFNPIEAERLSITTT
 PDAIQHIXPADHFVPGDISAFAFFIVAAALITPESDVTIHWGINSRPSGIIDIVERKMG
 GNIQLOFTTGAETPASIRIQYTPMLOPITIEGELVPAIDELPVALLCVTOAVGTST
 IKDAELKVKETNRIDTADMLNLGFELOPTNDGLIIPHSEFKTNATVDSLTDRHIG
 MMLAVASLSSEPVKIQPDANVNSPFGFLPKLLENEG"
 complement(4726. .5790)
 /gene="arob"
 complement(4726. .5790)
 /gene="arob"
 /note="ORFID:SA1298"
 /codon_start=1
 /transl_table=11
 /product="3-dehydroquinate synthase"
 /protein_id="BAB42558.1"
 /db_xref="GI:13701264"
 /translation="MKLQTPYPSNNTPIFEVHGAIDHISTYIDQFDSFILIDEHVNO
 YFADKFDIILSYENHVKVIIPAGEKTFEQYQETLEYILSHHVTNTAITAVGGGAT
 GDFAGFAATLRLGVHFIQVPTTILAHDSVGVKGINSKQKNILOGAFYRPTAVYD
 LDFLITLPPFOILSGYAEVYKHALNGESTTOEIQHFKDREILQSLMGMDKYIAGKI
 ETKLDIVVADEKEGVRKFLNLGHTGHAVENHKLAKCHAVMIGIIVQETIVANILEN
 SNHDIQHYINKLKLGYPLETIDIDFETIYQYMLSDKNDAQGVQWVLIHFGDIVV
 QHIDQTLQHRACGLKTYFK"
 complement(5816. .6982)
 /gene="arob"
 complement(5816. .6982)
 /gene="arob"
 /note="ORFID:SA1299"
 /codon_start=1
 /transl_table=11
 /product="chorismate synthase"
 /protein_id="BAB42559.1"
 /db_xref="GI:13701265"
 /translation="MRYLTSGESHGQPLTVIVGIPANLEIKVEDINKEMFKRGCGYG
 RGRMQIEKDTVEIVSGVRNGYTLGSPITWVYVNTDFTHWRKIMGAADPISSEENMK
 RTITKPRGHADLVGGHKNHRLRNVLERSSARETAARVAYGALCVLLQQLDIDYI
 SRVVEIGIKDKDFYDSETEFKANLDRNDVRVDDSTIAQMRDKIDEAKNGDSIGGVY
 QVYVENMPVGVSYHYDRKLDGKIAGQVYSINAFKGSFGEGKAAEKPGSEIQDEI
 LYNSEICYRGSNHILGLEGCMNGMPIIVNGVMKPIPTLYKPLNSVDINTKEDFKAT
 IERSDCAVPAASIVCEHVVAEIAKALLEEFSQNSHIEQLQOQIADRQLNVEP"
 complement(7368. .7568)
 /gene="SA1300"
 complement(7368. .7568)
 /gene="SA1300"
 /note="ORFID:SA1300"
 /codon_start=1
 /transl_table=11

		/product="hypothetical protein"			
		/protein_id="BAB42560.1"			
		/db_xref="GI:13701266"			
		/translation="MKNIISIIILGILMLPLKMLMELLYGAIFLORLPINPTIKIIFILTLI			
		YIFVIVRELIIIFUKSRNKA"			
gene		complement(7597..7722)			
		/gene="SAS045"			
CDS		complement(7597..7722)			
		/gene="SAS045"			
		/note="ORFID:SAS045"			
		/codon_start=1			
		/transl_table=11			
		/product="hypothetical protein"			
		/protein_id="BAB42561.1"			
		/db_xref="GI:13701267"			
		/translation="MWLVIMIIHKKARFLCCSKLLEKFSYFLVTRIKRGIFFIK"			
gene		complement(7777..8226)			
		/gene="ndk"			
CDS		complement(7777..8226)			
		/gene="ndk"			
		/note="ORFID:SA1301"			
		/codon_start=1			
		/transl_table=11			
		/product="nucleoside diphosphate kinase"			
		/protein_id="BAB42562.1"			
		/db_xref="GI:13701268"			
		/translation="WERTFLMIKPDVORNLEGEVSRERKRLKLVGKGLMQVPMEL			
		AETHYGEHQGPFFNDLITSPITSAPFAMVVEGEDAVNYSRHHIGSTNPSSEASPGSIR			
		GDLGLTVGRNIIHGSDLSLESAEREINLWENENETTSVSPRDALYE"			
gene		complement(8318..9277)			
		/gene="gerCC"			
CDS		complement(8318..9277)			
		/gene="gerCC"			
		/note="ORFID:SA1302"			
		/codon_start=1			
		/transl_table=11			
		/product="heptaprenyl diphosphate synthase component II"			
		/protein_id="BAB42563.1"			
		/db_xref="GI:13701269"			
		/translation="MAKLNMMNEIKKVEORLEKAIKSDSVLPQASLHLLSSGGKRV			
		PAFVILSSQFCKDEQTSQITQVAVALELHMATLVHDDVIDKSDRRKRLISKWD			
		99.0%; Score 1292.2; DB 1; Length 301550;			
		Best Local Similarity 99.4%; Pred. No. 1.3e-214;			
		Matches 1297; Conservative 0; Mismatches 8; Indels 0; Gaps 0;			
QY	1	ATGACTAAACCTATAGTACCTATTGTAGGTAGGCTTAATCTAGCTAAATCTACAAATTTT	60		
DB	13778	ATGACTAAACCTATAGTACCTATTGTAGGTAGGCTTAATCTAGCTAAATCTACAAATTTT	13719		
QY	61	AATAGAAATAGTTGGAGAACGTTTCGATTCTCAAGATACGCCAGGTGTACACCGAGAT	120		
DB	13718	AATAGAAATAGTTGGAGAACGTTTCGATTCTCAAGATACGCCAGGTGTACACCGAGAT	13659		
QY	121	CGTATTATCTTCAGGTAGTGGTTAACACATGATTTCAAATATTATTGATACAGGTGT	180		
DB	13658	CGTATTATCTTCAGGTAGTGGTTAACACATGATTTCAAATATTATTGATACAGGTGT	13599		
QY	181	ATTGAATTTGGTGATGCACCAATTCACAAACACAAATTAGAGCGGAGGAGAAATCGGCATA	240		
DB	13598	ATTGAATTTGGTGATGCACCAATTCACAAACACAAATTAGAGCGGAGGAGAAATCGGCATA	13539		
QY	241	GATGAAGCGGATGTTATTATTATTTTATGTTAACTGCGTGAAGGATTGACACAAACCGAT	300		
DB	13538	GATGAAGCGGATGTTATTATTATTTTATGTTAACTGCGTGAAGGATTGACACAAACCGAT	13479		
QY	301	GAATGGTGCCTCAAAATTTTATACAAATCTAAAAACCGGTGCTATTAGCGGTTAACAAA	360		
DB	13478	GAATGGTGCCTCAAAATTTTATACAAATCTAAAAACCGGTGCTATTAGCGGTTAACAAA	13419		
QY	361	GTAGATAATATCGAAATCGGTACAGAGCTGTATGATTTCTTATTATTAGGATTGGTGAA	420		
DB	13418	GTAGATAATATCGAAATCGGTACAGAGCTGTATGATTTCTTATTATTAGGATTGGTGAA	13359		
		Query Match 99.0%; Score 1292.2; DB 1; Length 301550;			
		Best Local Similarity 99.4%; Pred. No. 1.3e-214;			
		Matches 1297; Conservative 0; Mismatches 8; Indels 0; Gaps 0;			
QY	421	CCGTATCCAAATATACAGGGTCACATGCTTTAGGTCTTGTGTAGCTTGTGTAGATTCGATGTT	480		
DB	13358	CCGTATCCGATATACAGGGTCACATGCTTTAGGTCTTGTGTAGCTTGTGTAGATTCGATGTT	13299		
QY	481	TCTCATTTTGGTGAAGAGGAAGATCCTTATGATGAAGATACAAATTCGACTACTCATTCATT	540		
DB	13298	TCTCATTTTGGTGAAGAGGAAGATCCTTATGATGAAGATACAAATTCGACTACTCATTCATT	13239		
QY	541	ATTGGAGCACCACCTAGGTAAATCAAGTTTAGTAAATGCTATTTTAGGTGAAGATCGC	600		
DB	13238	ATTGGAGCACCACCTAGGTAAATCAAGTTTAGTAAATGCTATTTTAGGTGAAGATCGC	13179		
QY	601	GTTATCGTTTCTAATGTTGCAGGACACACGAGACGCTATTGATACAGAGTATAGTTAT	660		
DB	13178	GTTATCGTTTCTAATGTTGCAGGACACACGAGACGCTATTGATACAGAGTATAGTTAT	13119		
QY	661	GATGGACAAGATTATGTTTAAATCGATCTGCTGCTATGCTGCTTAAAGGAAAAGATATAT	720		
DB	13118	GATGGACAAGATTATGTTTAAATCGATCTGCTGCTATGCTGCTTAAAGGAAAAGATATAT	13059		
QY	721	GAATCAACTGAGAAATATTACGATTATTAAGAGCTTTTAAAGCGATTTGAACGTTCAAAATGTT	780		
DB	13058	GAATCAACTGAGAAATATTACGATTATTAAGAGCTTTTAAAGCGATTTGAACGTTCAAAATGTT	12999		
QY	781	GTTTGTAGTGGTTATTCATGCAGAACAGGCATTATTGAACAGATAAACGTTGTCAGGA	840		
DB	12998	GTTTGTAGTGGTTATTCATGCAGAACAGGCATTATTGAACAGATAAACGTTGTCAGGA	12939		
QY	841	TATGCACATGAACAGGTAAAGCAGTCGTGATGCTGCTGCTTAAATTAATGGGATACGTGCGAA	900		
DB	12938	TATGCACATGAACAGGTAAAGCAGTCGTGATGCTGCTGCTTAAATTAATGGGATACGTGCGAA	12879		
QY	901	AAAGATAGTAAAAACGATGAAGAAATTTGAAGATGAAGTACGTGAAGAAATTTCAATTTTAA	960		
DB	12878	AAAGATAGTAAAAACGATGAAGAAATTTGAAGATGAAGTACGTGAAGAAATTTCAATTTTAA	12819		
QY	961	GATTATGCACAAATTCCTTTGCTGCTGAAGAACCCACCAAGATACGTACATATTATTC	1020		
DB	12818	GATTATGCACAAATTCCTTTGCTGCTGAAGAACCCACCAAGATACGTACATATTATTC	12759		
QY	1021	CCTTACATTAATGAACCAAGTGAACCAATAAAAACCGTGTTCAAAGTTTCAACGTTTAAAT	1080		
DB	12758	CCTTACATTAATGAACCAAGTGAACCAATAAAAACCGTGTTCAAAGTTTCAACGTTTAAAT	12699		
QY	1081	GAAGTCGTTACTGATCAATTTCCATGAACCCCTACACCAACAGAGAGAGAGAGAGAGAG	1140		
DB	12698	GAAGTCGTTACTGATCAATTTCCATGAACCCCTACACCAACAGAGAGAGAGAGAGAGAG	12639		
QY	1141	AATGTCCTTTTATGCACACACAGTTGCTATAGAACCCACCGACATTTGTTGTTATTGTTAAAT	1200		
DB	12638	AATGTCCTTTTATGCACACACAGTTGCTATAGAACCCACCGACATTTGTTGTTATTGTTAAAT	12579		
QY	1201	GATGTAGAATTAATGCAATTTTCTTATTAACCGCTATTTAGAGAATCAATCCGTCGCGCT	1260		
DB	12578	GATGTAGAATTAATGCAATTTTCTTATTAACCGCTATTTAGAGAATCAATCCGTCGCGCT	12519		
QY	1261	TTTGGTGTGAGGTACACCAATTCATATTATAGCTCGAAAGAGA	1305		
DB	12518	TTTGGTGTGAGGTACACCAATTCATATTATAGCTCGAAAGAGA	12474		
		RESULT 2			
		AP003362/c			
		LOCUS			
		DEFINITION			
		Staphylococcus aureus subsp. aureus Mu50 genomic DNA, complete			
		sequence, section 5/9.			
		ACCESSION			
		AP003362 BA000017			
		VERSION			
		AP003362.2 GI:14247083			
		KEYWORDS			
		Staphylococcus aureus subsp. aureus Mu50			
		Staphylococcus aureus subsp. aureus Mu50			
		Bacteria; Firmicutes; Bacillales; Staphylococcus.			


```

/ gene="MW1341"
/ complement(278..619)
/ gene="MW1341"
/ note="ORFID:MW1341"
/ codon_start=1
/ transl_table=11
/ product="hypothetical protein"
/ protein_id="BAB95206.1"
/ db_xref="GI:21204510"
/ translation="MTNITKEVDFNLEQIDVPAKNKTLGSSKAPVLDVHSHIIDY
PKQVNDITGNIDPKLNQIPVVPVMEKERYDIAMIERKHFGYQMKTFKTELKMAA
SYGTRLNKQV"
/ complement(624..1283)
/ gene="nth"
/ complement(624..1283)
/ gene="nth"
/ note="ORFID:MW1342"
/ codon_start=1
/ transl_table=11
/ product="endonuclease-like protein"
/ protein_id="BAB95207.1"
/ db_xref="GI:21204511"
/ translation="MVSRAKALEMIDVIANMFPDACELKNDNFELTIYAVLLSAQCT
DVLNVATTELEFKYKTPEDYLAVSDEELMINDIRSIGLYRNKAKNIRKLCQSLIDQFN
GEIPTHKELESLAGVGRKTANNVMSVAFDEPSLAVDTHVERVSKRIGINRWKQNVRO
VEDRLCSVIPRQNRNRRHQLIFFGRYHCLARKPKDCICPLEDCREGQRYRASLKE
A"
/ complement(1273..1959)
/ gene="MW1343"
/ complement(1273..1959)
/ gene="MW1343"
/ note="ORFID:MW1343"
/ product="hypothetical protein, similar to chromosome replication
initiation protein dnaD"
/ codon_start=1
/ transl_table=11
/ protein_id="BAB95208.1"
/ db_xref="GI:21204512"
/ translation="MDKYOLKARPVVIRRELDHYSDIGLDEODLVLLKLIYASETS
NKQPSLELLQKGTMPDRDITHVIONLQRELELOVQKDEGRFTYMLNDDPFEKL
SHILKQSTKETEONSKFKOLFVLEDTFARPLSYEITLQWIDVQKHUSALIQ
AALDEANSKLKLSFKYMDRILLNWKNNKVIDDSRKIREKFNKMTHTVTKVTPFD
WLANGLDK"
/ complement(1287..3579)
/ gene="asn"
/ complement(1287..3579)
/ gene="asn"
/ note="ORFID:MW1344"
/ codon_start=1
/ transl_table=11
/ product="asparaginyl-tRNA synthetase"
/ protein_id="BAB95209.1"
/ db_xref="GI:21204513"
/ translation="MKTITKQKHLNDQVITGAWLTNRKSSGKIAFLQRLDGTGPMQ
GVVXSEVDEEYFKLAKETQESSLYVTGTITEDNRSDLGEMQVKSIEVISAHDYP
ITPKNGTEFLMDRHLRLRSKQKOHAVMKIRNEVIRATYFENKDGFTKVDVPSILTAS
APEGTSELEFTRYFQDAPFLSQSGOLYLEAAMAHGKVFSPGPTFRAEKSSTRHLIE
FMIEGMAFTNHAESLEIQEYVTHVVKSVLENCKLEKILERTDTSKLEKVAIPPR
ISYDDAIETPLKSGEDDIEMGEDFCAPHETAIANHYDLVPVITNYPTKIKPFYKQNP
ENETVLCADLIAPCEYGEIIGSERVDDLELLEQVKEHCLDEEAYSYYLDLRRYGS
VPHCGGLGELTEVAWISGVENHRETAPFPRLNRLYP"
/ complement(1901..6594)
/ gene="ding"
/ complement(1901..6594)
/ gene="ding"
/ note="ORFID:MW1345"
/ codon_start=1
/ transl_table=11
/ product="probable ATP-dependent DNA helicase ding"
/ protein_id="BAB95210.1"
/ db_xref="GI:21204514"
/ translation="MGMTAVVQLETTGNGQLDFDDIIQIGITFVRNQNIIQTHYSMI
RTMLEIPPIQALTSTERNMLQQQAPYFNQVQAEIYDKIKOCIPVAHNVDPLNFIKKA

```

```

/ gene="MW1346"
/ complement(6618..7589)
/ gene="MW1346"
/ note="ORFID:MW1346"
/ codon_start=1
/ transl_table=11
/ product="truncated hypothetical protein"
/ protein_id="BAB95211.1"
/ db_xref="GI:21204515"
/ translation="MSKYSQVQLQLYKKNPNYISQSTAESLNTSRVAKKVIDOLK
LEGCKIDSVNKHGHLQLOPDINYOQIIDQVTKSSALDFSEVDSIDSTOLAKKSL
VGNOSSEFLLSDQTKGRGFRNHNSSSGCGLMHSVLRPNVAFSHISKFNLTALG
IRDAIQHFSQDEVKYKPNNDIYIDNGYCGFLTEVANNNGTEATIGIGINLTQOLE
NFDESIRHRATSVQLHDKNLRDQFLERLQAEIERKYNQFLPFPFSIREEYTAASN
IWNRTLLFTENDQKFGQAIDLDYDGYLIVRDAGESHRLISADIDF"
/ complement(7576..8778)
/ gene="MW1347"
/ complement(7576..8778)
/ gene="MW1347"
/ note="ORFID:MW1347"
/ product="hypothetical protein, similar to poly(A) polymerase"
/ codon_start=1
/ transl_table=11
/ protein_id="BAB95212.1"
/ db_xref="GI:21204516"
/ translation="MKSLSFEQARPILEDIQNGFEAYVGVGSVRYVNGRNHIDDI
TTSAPDEIESIFSHTPVKGHTINNVFNDENYEVTTFAEEDYVDHRRPQSGVTFV
RDLYEDLQRDEFTMAINADMTAYKLYDQDQDINNRRIIRTGTAEERFQEDALRMI
KRLRFQSLQSDIATFEAMRIQMAIDNFIERSIERVIELTKMRGINVKSFNHLKS
LIAFNMPYFEHLOMNOINVTEAIDLELLIAVSVKFOINYSKPLKLSNRQVADINO
YIQIMNALPSIITKEOLKNFVYDTHLIKNNVAAVADIKANDIQGHPLVNLQITID
ETLHRLPMHNRKDMVNGSVLMAHLNAKSGPLKDVLRQIEIAITVGRVSNTEELK
WVDNHNKI"
/ complement(8783..9925)
/ gene="MW1348"
/ complement(8783..9925)
/ gene="MW1348"
/ note="ORFID:MW1348"
/ product="hypothetical protein, similar to lipopolysaccharide
biosynthesis-related pr homolog"
/ codon_start=1
/ transl_table=11
/ protein_id="BAB95213.1"
/ db_xref="GI:21204517"
/ translation="MKGITCTPSMGCGSIIATELGIKLAERGHEVHTSINIPFIR
KPLPMHIFHVEVQNAVFOYPPYDITLSTKIAEVIKGYDOLLHMHYAVPHATCGIL
ARENSGDKIKIMTTLHGDTITVLGYDHSIQGAIKRGIKSDIYTSVSKLAEQTHEII
ETNKEIPIYFVRENEFTKHTALKSQGIAPDEKVLINVSFPROVKRITDIETFF
AKVREKIPSKLILLGDGPBLVPMRQLTKELNVEEDVLEKQDCVSFFYQLSOLVLL
SEKESFGLTLEAMKTGVPICSGNAGGKEVIKHGCTGTVVDVGDCDSADYAIRLLE
DKVLYNKLQKNMLADIAERFGSELITDQVEYYVYONKMLNEHNKSGE"
/ complement(10167..10484)
/ gene="MW1349"
/ complement(10167..10484)
/ gene="MW1349"
/ note="ORFID:MW1349"
/ codon_start=1
/ transl_table=11

```

/product="conserved hypothetical protein"	
/db_xref="GI:21204518"	
/translation="MSKVKQREVEYIGORFQVSPFLANLRLTEVCELAREIN	
HTYGEKKKKSEENATKAEIGDNLFLVLLCLANSMDIDTFESFNETMEKFNTRDKRF	
ERK	
gene	complement(10820..11518)
CDS	/gene="MW1350"
	complement(10820..11518)
	/gene="MW1350"
	/note="ORFID:MW1350"
/codon_start=1	
/transl_table=11	
/product="conserved hypothetical protein"	
/protein_id="BAB95215.1"	
/db_xref="GI:21204519"	
/translation="MSSIFTHIYFVLWVLPVLAHQHKVSKYKYSQVRSFSGTKGR	
EVALLEIHANGIYDEVYKGEFLDHYDPKKVWSLSNPANDRPSVAGTAIAAHEVG	
HAIQDHQYVFLFRALVFPVNLGSSLSYHIMLGIILTAIGSAGFTALWIGALM	
Query Match	
Best Local Similarity 98.9%; Score 1290.6; DB 1; Length 333750;	
Matches 1296; Conservativeness 0; Mismatches 9; Indels 0; Gaps 0;	
Qy	1 ATGACTAACTATAGTACGCTATGTTAGGTAGCCCTAATCTAGGTAAATCTCAATTTT 60
Db	
Qy	24351 ATGACTAACTATAGTACGCTATGTTAGGTAGCCCTAATCTAGGTAAATCTCAATTTT 24292
Db	
Qy	61 AATAGATAGTTGGAGACGTTGTTTCGATTGTTGAAGATACGCCAGGTCTAACACGAGAT 120
Db	
Qy	24291 AATAGATAGTTGGAGACGTTGTTTCGATTGTTGAAGATACGCCAGGTCTAACACGAGAT 24232
Db	
Qy	121 CGTATTATCTTCACGTGAGTGGTTAACACATGATTTCATATATTATCATACAGTGGT 180
Db	
Qy	24231 CGTATTATCTTCACGTGAGTGGTTAACACATGATTTCATATATTATCATACAGTGGT 24172
Db	
Qy	181 ATTGAATTTGTTGATGACCAATCCAAACACAAATTAGACGCGACGACAAATCCCAT 240
Db	
Qy	24171 ATTGAATTTGTTGATGACCAATCCAAACACAAATTAGACGCGACGACAAATCCCAT 24112
Db	
Qy	241 GATGAGCGGATGTTATTATTTTATGTTTAACTGCGTGAAGGATTGACACAAACGAT 300
Db	
Qy	24111 GATGAGCGGATGTTATTATTTTATGTTTAACTGCGTGAAGGATTGACACAAACGAT 24052
Db	
Qy	301 GAAATGGTGGCTCAAAATTTATACAAATCTAAAAACCGTCTGTTAGCGTTAACAAA 360
Db	
Qy	24051 GAAATGGTGGCTCAAAATTTATACAAATCTAAAAACCGTCTGTTAGCGTTAACAAA 23992
Db	
Qy	361 GTAGATAATATGGAATCGGTACAGACGTTGATGTTCTTATTCATTAGGATTGGTGAA 420
Db	
Qy	23991 GTAGATAATATGGAATCGGTACAGACGTTGATGTTCTTATTCATTAGGATTGGTGAA 23932
Db	
Qy	421 CCGTATCCAATATCAGGTCACATGGTTTGGTCTTGGTCTGTTAGTACGTTGTT 480
Db	
Qy	23931 CCGTATCCAATATCAGGTCACATGGTTTGGTCTTGGTCTGTTAGTACGTTGTT 23872
Db	
Qy	481 TCTCATTTTGGTGAAGAGAGATCCCTTATGATGAAGATACAAATTCGACTATCCATT 540
Db	
Qy	23871 TCTCATTTTGGTGAAGAGAGATCCCTTATGATGAAGATACAAATTCGACTATCCATT 23812
Db	
Qy	541 ATTGACGACCAACGTTAGGTAATCAAGTTTATGTAATGCTATTTTAGGTGAAGATCGC 600
Db	
Qy	23811 ATTGACGACCAACGTTAGGTAATCAAGTTTATGTAATGCTATTTTAGGTGAAGATCGC 23752
Db	
Qy	601 GTTATCGTTTCTTAATGTTGACGAGCAACGAGACGCTATTGATACAGATATAGTAT 660
Db	
Qy	23751 GTTATCGTTTCTTAATGTTGACGAGCAACGAGACGCTATTGATACAGATATAGTAT 23692
Db	
Qy	661 GATGACAGATATGTTTTTATCGACTCTGCTGCTATGCTGTAATGCGTAAAAAGGAAAGTAT 720
Db	
Qy	23691 GATGACAGATATGTTTTTATCGACTCTGCTGCTATGCTGTAATGCGTAAAAAGGAAAGTAT 23632
Db	
Qy	721 GAATCACTGAGAAATATTCAGTATTAAAGACCTTTAAAGCGATTGAACGTTCAATGTT 780
Db	

Db	23631	GAATCACTGAGAAATATTCAGTATTAAAGACCTTTAAAGCGATTGAACGTTCAATGTT	23572
Qy	781	GTTTATAGTGGTTATTGATGCGAGAACAGGCAATTATTGAACAGATAAACGTTGTCAGGA	840
Db	23571	GTTTATAGTGGTCATAGATGCGAGAACAGGCAATCATTGAACAGATAAACGTTGTCAGGA	23512
Qy	841	TATGCACATGAACAAGGTTAAAGCAGTCGTGATTGTCGTTAAATAAATAAATGGGATACTGTGNA	900
Db	23511	TATGCACATGAACAAGGTTAAAGCAGTCGTGATTGTCGTTAAATAAATAAATGGGATACTGTGNA	23452
Qy	901	AAAGATAGTAAACGATGAAGAAATTTGAAGATGAAGTACGTAAAGAAATTTCAATTTTFA	960
Db	23451	AAAGATAGTAAACGATGAAGAAATTTGAAGATGAAGTACGTAAAGAAATTTCAATTTTFA	23392
Qy	961	GATTATGACCAATTTGCTTCTGCTAAAGACGACAGATTTAGTACATATTTC	1020
Db	23391	GATTATGACCAATTTGCTTCTGCTAAAGACGACAGATTTAGTACATATTTC	23332
Qy	1021	CCTTACATTAATGAAGCAAGTGAACCAATTAACCAAGCTGTTCAAGTTCAACTTTTAAAT	1080
Db	23331	CCTTACATTAATGAAGCAAGTGAACCAATTAACCAAGCTGTTCAAGTTCAACTTTTAAAT	23272
Qy	1081	GAAGTGTGTTACTGATGCAATTTCCATGAACCTTACACCAACACACAAAGGTAGCGTTG	1140
Db	23271	GAAGTGTGTTACTGATGCAATTTCCATGAACCTTACACCAACACACAAAGGTAGCGTTG	23212
Qy	1141	AATGCTTTTATGCAACACAGTTGCTATAGAACCCAGACATTTGTTGTTATTTAT	1200
Db	23211	AATGCTTTTATGCAACACAGTTGCTATAGAACCCAGACATTTGTTGTTATTTAT	23152
Qy	1201	GATGATAGATTAATGCAATTTTCTTATAACGCTATTATTAGAGATCAATCCGTCGCT	1260
Db	23151	GATGATAGATTAATGCAATTTTCTTATAACGCTATTATTAGAGATCAATCCGTCGCT	23092
Qy	1261	TTTGGTTTGAAGGTACACCAATTCATATTATAGCTCGAAGAGA	1305
Db	23091	TTTGGTTTGAAGGTACACCAATTCATATTATAGCTCGAAGAGA	23047

RESULT 5	AF270032/c	3269 bp	DNA	linear	BCT 01-AUG-2000
LOCUS	Staphylococcus epidermidis strain SRI clone step.1042f07 genomic				
DEFINITION	sequence.				
ACCESSION	AF270032				
VERSION	AF270032.1	GI:9623936			
KEYWORDS	Staphylococcus epidermidis				
SOURCE	Staphylococcus epidermidis				
ORGANISM	Bacteria; Firmicutes; Bacillales; Staphylococcus.				
REFERENCE	1 (bases 1 to 3269)				
AUTHORS	Kimberly, W.J., Taylor, J. David., Nelsen, A.J., Godlevski, M.M., Rubino, M.A., Nelson, F.J., Rivers, P.R., Torruella-Miller, I., Lisenbee, S., Ashanti, C., Altschuler, G., Mamo, L., Shepherd, N.S., Fuchs, R., Fleming, T., Guan, X., Du, L., Cain, D.H., Miller, G.S. and Furdon, P.J.				
TITLE	Transposon-mediated sequencing of the Staphylococcus epidermidis genome				
JOURNAL	unpublished				
REFERENCE	2 (bases 1 to 3269)				
AUTHORS	Taylor, J. David., Kimberly, W.J., Nelsen, A.J., Godlevski, M.M., Rubino, M.A., Nelson, F.J., Rivers, P.R., Torruella-Miller, I., Lisenbee, S., Ashanti, C., Altschuler, G., Mamo, L., Shepherd, N.S., Fuchs, R., Fleming, T., Guan, X., Du, L., Cain, D.H., Miller, G.S. and Furdon, P.J.				
TITLE	Direct Submission				
JOURNAL	Submitted (22-MAY-2000) Departments of Genomic Sciences and Bioinformatics, Genetics Directorate, Glaxo Wellcome, Inc., 5 Moore Drive, Research Triangle Park, North Carolina 27709-3398, USA				
FEATURES	Location/Qualifiers				
source	1. 3269				
	/organism="Staphylococcus epidermidis"				

Db	1904	AAAGTAGTAAAGCAAAATGAAATAATTTCACTGATGATGTTAGAAATGAATTTCAATTTT	184
Qy	961	GATTATGACAAAATTTGCTTTTGTGCTGCTAAAGAACGCACAAAGATTACGTACATATTTC	1020
Db	1844	GATTATGCTCAAAATCCGCTTCTGATACGAAAGAGGCTTAAGATTTAAACAAATTTATTC	1785
Qy	1021	CCTTACATTAATGAAGCAAGTGAAACCCATAAACAGCTGTTCACAGCTTCAACTTTAAAT	1080
Db	1784	CCTTATATCAATCAAGCCAGTGAAATCATAAAGGCTGTCACAAAGCTTCTACACATAAT	1725
Qy	1081	GAAGTCGTTACTGATGCAATTTCCATGAACCCCTACACCAACAGACAAAGGTACAGCTTTC	1140
Db	1724	GAAGTGTGTACTGATGCTCAATCTCTATGAATCCAAACCTCTACTGACAAAGGTACAGACCT	1565
Qy	1141	AATGTCTTTTATGCAACACAAAGTTGCTATAGAACCACCGACATTTTGTGTATTGTTAAAT	1200
Db	1664	AATGTATTCTATACAACTCAGTGTGCAATTCGAACCCAGACATTTGTAGTATTGTCAAT	1605
Qy	1201	GATGTAGAATTAATGCATTTTCTTATTAACGCTATTAGAGAAATCAATCCGTCGCGCT	1360
Db	1604	GATGTGAATTAATGCATTTTCTTATAGGAGATATTATAGAAATCAATACGTAAATGCT	1345
Qy	1261	TTTGGTTTTGAAGGTACACCAATTCATATTATAGCTCGAAAGAGA	1305
Db	1544	TTTGGTTTTGAAGGAACACCTATTTCATATTATTCGAAGAAAAGA	1500
RESULT 6			
LOCUS	AX145350/c	3269 bp	DNA
DEFINITION	Sequence 4072 from Patent WO0134809.	linear	PAT 31-MAY-2001
ACCESSION	AX145350		
VERSION	AX145350.1	GI:14283915	
KEYWORDS	synthetic construct		
SOURCE	synthetic construct		
ORGANISM	artificial sequences.		
REFERENCE	1		
AUTHORS	Kimmerly, W.J.		
TITLE	Staphylococcus epidermidis nucleic acids and proteins		
JOURNAL	Patent: WO 0134809-A 4072 17-MAY-2001;		
FEATURES	GLAXO GROUP LIMITED (GB)		
Source	Location/Qualifiers		
	1..3269		
	/organism="synthetic construct"		
	/mol_type="genomic DNA"		
	/db_xref="taxon:32630"		
	/note="synthetic nucleic acid sequence"		
BASE COUNT	986 a 616 c 451 g 1216 t		
ORIGIN			
Qy	1	ATGACTAAACCTATAGTAGCTATGCTAGCTAGGCTCAATGCTAGCTAAATCTACAAATTTT	60
Db	2804	ATGACTAAACCCGATAGTACCAATTTGAGGAACCAACCAATGTAGGAATCTACAAATTTT	2745
Qy	61	AATAGAATAGTTGGAGAACGTTTTCGATTCTTTGAAGATACGCCAGGTGTAAACACGAGAT	130
Db	2744	AATAGAGTTGTGGCGCAACGATATATCTATCTAGTAAGATACGCCAGCGCTTACTCGTAT	2685
Qy	121	CGTATTATTCTTCACAGTGAGTGGTTAAACACATGATTTCAATATTATTATGATACAGGTGT	180
Db	2684	CGTATTTTATTCATCTGCGCAGTGGTTAACTCAATTTTAACATCATTTGATACAGGTGT	2625
Qy	181	ATTGAATTTGGTGTACCACTTCCAAACACAAATTTAGCGGACGAGCAAAATCGCCATA	240
Db	2624	ATTGAATTTGGAGATGCTCCTTTTCAAACCTCAATCCGTCACAGGCGAAGATAGCAAT	2365
Qy	241	GATGAACGCGATGTTATTATTTTATGTTTAACTGCGTGAAGGATTGACACAAAGCGAT	300
Db	2564	GAAGAAGCAGATGTCATCATTTTATGCTCAATGTTAGAGAAGGACTTACACAAAGTGAC	2505
Qy	301	GAATGCTGCTCAAAATTTTATACAAATCTTAAACCCGCTGATTAGCGGTTTAAACAA	360
Db	2504	GAATGCTGCGCAAACTGTTTAACTTAAGAACTGTTGTTATGCTGTGAATATA	2445
Qy	361	GTAGATATATGGAATCGGTACAGACGCTGATGATTTCTATTCTATTAGGATTTGGTGAA	420
Db	2444	GTGTGATATCTTGAATTCGCTATATCATATCATGATTTCTATTCTTTAGGCTTTGGAGAT	2385
Qy	421	CGGTATCCAAATATCAGGTCACATGGTTTAGTCTTGTGACTTTGTTAGATGCAAGTTGTT	480
Db	2384	CCATATCCTATTCTGTTTACATGGATTAGGACTTTGGAGATTTGCTAGATGCAAGTTGTT	2325
Qy	481	TCTCATTTTGGTGAAGAGGAAGACATCCTTATGATGAAGATACAATTCGACTATCCATT	540
Db	2324	GAAACTTTTAAAGAAATCAGAAATCCTTATGACGAAGATAGATACGCTTTTCTATC	2265
Qy	541	ATTGACGACCAACAGCTAGTAAATCAAGTTTAGTAAATGCTATTTTAGGTGAAGATCGC	600
Db	2264	ATCGGTAGACCTTAATGTTGGTAAATCTAGCTTGGTCAATGCTATTTTAGGCGAAGAACGT	2205
Qy	601	GTTATCGTTTCTTAATGTTGCGAGGACACGAGACGCTATTGATACAGATATAGTATTAT	660
Db	2204	GTTATTGTCTTAATGTTGCTGCTACAACTCGAGATGCTATTGATACCGAGTACTCTTAT	2145
Qy	661	GATGGACAGATTATGTTTAAATCCATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	720
Db	2144	GATGGACAGATTATGTTTAAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	2085
Qy	721	GAATCACTGAGAAATATTTCAGTATTAAAGAGCTTTTAAAGCGATTGAACGTTCAATGTT	780
Db	2084	GAATCGACTGAAATAATTTCTGTTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	2025
Qy	781	GTTTATGCTGTTATTGATGCGAAGCAAGGCTATTGTAACAGATTAACAGCTGTTGTCAGGA	840
Db	2024	GTTATTAGTATTATCGATGCTGAACAAAGTATAATTGTAACAGATTAACAGCTGTTGCTGCT	1965
Qy	841	TATGCACATGAACAGGTAAGCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	900
Db	1964	TATGCACATGAGAGGTAAGCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	1905
Qy	901	AAAGATAGTAAACCGATGAAGAAATTTGAAGATGAAGTACGCTAAAGAAATTTCAATTTT	960

VKLTKMSIDELCPVNLNAIEQFEELNERYTFLNEQRTDLREAKETLEQIIHENDKEVE
GRFKTTHAVQDHFTTVFKQFGGQAEHLRLTEDDYLISAGVDIIVQPPKKLQHLULL
SGERALSIALFAILVAGAPFVILDEBAALDEANVIRIAYQYNELSTETQFIVI
THKRGTFMSDRLYGVTHQESGVSKLVSVNLWTIDEVKNBEQA

gene

4891..6117

CDS

/locus_tag="SE0910"

4891..6117

/locus_tag="SE0910"

/codon_start=1

/transl_table=11

/product="signal recognition particle"

/protein_id="AA004513.1"

/db_xref="GI:27315373"

EOVKPKKKKLESEADDFDEGLISIEDFEETEAQKIGAKFKAGLEKSRQNFQOLNNL
IARYKVDYDFEFALEEMLTADVGFNTVMKLTDELATQAORRNQIQTEDRLRIVIEK
IYIYHQEDDSZAMNIEDGRNLVILMGVNGVGTITTKLAYRYOECKKVMLAG
DYFRAGAIQQLNMGVERGVVEVVSQNEGSDPAAYVDAINAAKNKGVDIILICITACRL
QNKSNLQELDKMKRVINRAIPDAPHEALCLDATTGONALSOARSKFVTVNSGIVL
TKLDGTAKGIVLAIRNELHIPVKYVGLGERKMDLQPPNPESVYVGLFADMIQNEDI
PREISRSVSVEEGN

gene

6119..6451

CDS

/locus_tag="SE0911"

6119..6451

/locus_tag="SE0911"

/codon_start=1

/transl_table=11

/product="conserved hypothetical protein"

/protein_id="AA004508.1"

/db_xref="GI:27315374"

/translation="MGQNDLVKTLRMNVLDFYOSLTLTKQKNLYEFFYLQDYSLSSEI
ADTFEVSRAVYDINRTGDLVEDYESKRLYQFRRRELYNLKQSLNQPELLQY
ITOLEELE

gene

6482..7849

CDS

/locus_tag="SE0912"

6482..7849

/locus_tag="SE0912"

/codon_start=1

/transl_table=11

/product="fifh protein"

/protein_id="AA004509.1"

/db_xref="GI:27315375"

/translation="MAFELSDRLQATMQKMRGKGVTEADITKTMREYRALLLEADV
NFKVKEFVTVSERALSDVQSLTPCGQVIKIVQBELSLMGMENTSILKMANKEPT
VVMVQLOGAGTTAGKLALLMRKNKPLLVAGDIYRPAADIQOTVYQKQIDIPV
YSEQGVQPOQIVENALKHAKKEHLDFYIIDTAGRLHIDEALNELQEVKEISKDEI
MLVVDANTQDAVAVQAQFDEQLDVSGLTKLDGDTRGGAALSIRSVTKPKFVGM
SEKLDGELFHPERMASRILGMDVLSLIEKAOQDVQERAKDLKMRSSFTLDFE
LEQLDQVKNLGPLDDIMKIPCMNMKGKGLKLNMSDKQIDHIIKALIOSMTPKERSNPE
TLNVSRRKRIAGSGRSLOEVNRLKQFNDKMKMKKQFTGGGKKKAKRQMQNMLKG
MNLPP"

gene

8281..8556

CDS

/locus_tag="SE0913"

8281..8556

/locus_tag="SE0913"

/codon_start=1

/transl_table=11

/product="30S ribosomal protein S16"

/protein_id="AA004510.1"

/db_xref="GI:27315376"

/translation="MAVKIRLTRGSKRNPFIYRIWADARSPDRGRIIEQLGTYNPAH
VNAPEVKRIDEALAWLHDGAKPTUTVHNILSRGILKKFDEQKNK"

gene

8683..9186

CDS

/locus_tag="SE0914"

8683..9186

/locus_tag="SE0914"

/codon_start=1

/transl_table=11

/product="rimM protein"

/protein_id="AA004511.1"

/db_xref="GI:27315377"

/translation="MEVEVGQIVNTHGKGVKVSNSDFTETRFQEGOLLVKVKNHT
EIVTIVASTIRKHGFHMLRFGGINNINIEHLKGDYIYQERHODIELGEHEYYISDI

IGCTVFKDDDTPIGRVINIFETGANDVWVWKEKEYLIPYIEDVVKOIDIETRTKIT
PMEGLLD

gene

9186..9923

CDS

/locus_tag="SE0915"

9186..9923

/locus_tag="SE0915"

/codon_start=1

/transl_table=11

/product="tRNA (guanine-N1)-methyltransferase"

/protein_id="AA004512.1"

/db_xref="GI:27315378"

/translation="MKIDYTLFPPEMFEGVLNHSILKRAQDKGIMVNTINFRDYSIN
KHNQDDYPCGGQGMVLKPEPVFNAMEDINNHETRVILMCPQGRPFTQIAQELSE
AKHVFCTGHEGYEDIRKHLVDTDEISMGDYVLTGGELPAMTMTDAIVRLIQGLGN
QASHQDQSDGLLEPPQYTRPREYKNMSVPEVLISGNHAHIDQWRHEQKLIRTYEKR
PDLLQYPTLEKOREILETYKKKLAND

gene

10029..10379

CDS

/locus_tag="SE0916"

10029..10379

/locus_tag="SE0916"

/codon_start=1

/transl_table=11

/product="50S ribosomal protein L19"

/protein_id="AA004513.1"

/db_xref="GI:27315379"

/translation="MSNHKLIKAVTKSOLRTDLPTFRSGDTRLRVHVRIVEGSRERIOV
PEGVVIKRGGISETFTVRKISSGVGVERTPLPTPIKIEVKRRGKVRRAKLYL
RSLRGKAARIQIEIR

gene

complement(10910..13513)

CDS

/locus_tag="SE0917"

complement(10910..13513)

/locus_tag="SE0917"

/codon_start=1

/transl_table=11

/product="conserved hypothetical protein"

/protein_id="AA004514.1"

/db_xref="GI:27315380"

/translation="MIKNWNPKIRFTLIVLVSILSFLYGLPPLYQFTKGIIVFSG
SGDGEFQMPFOMHLYHEFTSLKGYDASFGGDDYKSLAYYSLSPMLWNESMIM
ILEQITVNPNDISFNTIQLINAYRVITFESFYLSYLRKLPAPMFATILYGM
STVVTITFWSFTGNLLMLPMSLWATERFERKIKGWFIAIATLFSNFTSYTE
AIVGFIYTRFAPHEKDIIVNMOKLYILVCATULSVLSIYGLYTVGVSSFLDNDRT
ONPNRFTFTNFETNIFADGVYITISFIALICEFLKQHYHYKLFALFATWIL
LIGFSQMFDSAPNGFSLPQRMVYFALSTIALFQHLSEISEIKYFTFVAIPV
IYGFITALSERSVKMFVALIILVIFIKYKSLTRTSMVLLVLYFLAQOVLMT
NDRKTIETPYQTTIKTINDSSYRSPVLNKKIKYHQSSDTPDLRLDYFVYALNSPL
IYHNGTSLYSSIFDGLIKYDDOTLOINMPVDKNSYRVLNNRANLMSLWDVODRLR
HPDDLNNPYCFKKKELITDKKQDWHSYNTINYPASHITNTKYDARKLKSPLDQDAM
LKGVNLHNSOANDTFKPNLLSHAKONLHANNHDSKHLKVKOHGCVTLNLPRI
VKYKQNYIEMDVLLSPDKHKGVNEYSQERNLSYKYRFPVSPVTRAKASNQLN
IKMSGYIRFKRYAGTYGENTQTLKASQOLQPVKVKERNNGFTTIKKKKEHGYLVLPN
VYAGHHAMANGKPLKVQOQNGIMTTPVKEGQTKIKLSYTPPYFYLLITVSCIGIIL
SILFTHYVRRK

gene

complement(13506..15779)

Query Match

70.5% Score 919.4; DB 1; Length 300892;

Best Local Similarity 81.5% Pred. No. 4.1e-150;

Matches 1064; Conservative 0; Mismatches 241; Indels 0; Gaps 0;

Qy

.1 ATGACTAAACCTATAGTAGCTATTGTAGTAGGCCTAATGTAGGTAAATCTACAATTTT 60

Db

297184 ATGACTAAACCGATATAGCAATTTGAGGAAACAAATGTAGGAAATCTACAATTTT 297125

Qy

61 AATAAGATAGTTGAGAACCGTGTTCGATTTGTAAGATACCCAGGTTAACACAGAGAT 120

Db

297124 AATAGAGTTGTGCGCGAAGCTATATCTATCGTAGAAGATACGCCAGCGTTACTCGTAT 297065

Qy

121 CGTATTATTCTTCAGGTGAGTGGTTAACACATGATTTCATATATTATTGATACAGGTGCT 180

Db

297064 CGTATTATTTCATCTGCGGAGTGGTTAACTCATGAAATTTAACATCATTTGATACAGGTGCT 297005

Qy

181 ATTGAATTTGTTGATGACCAATTCACACAAATTTAGAGCCAGCGAATAATCCCAT 240

Db 297004 ATTGAATGGAGATGCTCTCTTTTCAAACTCAATCCGTGCACAGCAGAAATAGCAATT 296945
QY 241 GATGAAGCGGATGTTATTTATGTTTAACTGCTGTAAGGATTGACACAAAGCCAT 300
Db 296944 GAAGAGCGAGATGTCATCATTTTATGTTGTCATGTTAGAGAGGAGCTTACACAAAGTGC 296885
QY 301 GAATGTCGCTCAAAATTTTATCAAAATTTTAAACCAACCGTTCGTTATAGCGGTTAACAAA 360
Db 296884 GAATGTCGCGACAAATGCTTTTAAATCTTAAGAAACCTGTTGTTATAGCTGTGTAATAA 296825
QY 361 GTAGATAATATGGAATCGGTACAGACGCTGATGTTCTTATTTATTTAGGATTTGGTGAA 420
Db 296824 GTTGATAATCTGGAATCGGTAAATGATATCTATGTTCTTTAGGCTTTGGAGAT 296765
QY 421 CCGTATCAATATCAGGCTCACATGTTTGGTCTTGGTACTGTTAGATGACGTTCTT 480
Db 296764 CCATATCTTATTTCTGTTTCATGTTAGGACTTTGGAGATTTGCTAGTACGTTGTT 296705
QY 481 TCTCATTTTGTGAAGAGGAGAGATCTTTATGATGAAGATACAAATTCGACTATCCATT 540
Db 296704 GAAACCTTTAATAAGAAATCAGAAGATCTTTATGACGAAGATACGATCTCTTCTATC 296645
QY 541 ATTGACGACCAACCAAGCTAGGTAAATCAAGTTTAAATGCTTATTTAGGTGAAGATCGC 600
Db 296644 ATCGTAGACCTTAATGTTGTTAAATCTAGCTTGGTCAATGCTATTTAGGCGAAGAGCT 296585
QY 601 GTTATCGTTTCTAATGTTGACGAGACCAAGAGCGCTTATTTGATACAGAGTATAGTTAT 660
Db 296584 GTTATGTTGCTAATGTTGTTGTTACAACTCGAGATGCTTATGATACCGAGTACTTAT 296525
QY 661 GATGACAAAGATTTATGTTTAAATCGATCTGCTGTTATCGCTTAAAGAGGAAAGATATAT 720
Db 296524 GATGACAAAGATTTATGTTTAAATCGATCTGCTGTTATCGCTTAAAGAGGAAAGATATAT 296465
QY 721 GAATCACTGAGAAATTTATGATTAAGAGCTTTTAAAGCGATTTAAAGCTTCAATGTT 780
Db 296464 CAATCGATGAAATTTATGTTTAAATCGATCTGCTTAAAGCGATTTAAAGCTTCAATGTT 296405
QY 781 GTTTAGTGGTTTATGATGACGAGCAAGGCTTATTTGAACAGATTAACGCTGTCACAGA 840
Db 296404 GTATTAGTGGTTTATGATGACGAGCAAGGCTTATTTGAACAGATTAACGCTGTCACAGA 296345
QY 841 TATGACATGAACAGGTAAGGACGCTGCTGTTATGCTGTTAAATTAAGGATGCTGTAAG 900
Db 296344 TATGACATGAAGGTAAGGACGCTGCTGTTATGCTGTTAAATTAAGGATGCTGTAAG 296285
QY 901 AAAGATGTAACAGGTAAGGACGCTGCTGTTATGCTGTTAAATTAAGGATGCTGTAAG 296225
Db 296284 AAAGATGTAACAGGTAAGGACGCTGCTGTTATGCTGTTAAATTAAGGATGCTGTAAG 296225
QY 961 GATTATGACAAATTTGCTTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1020
Db 296224 GATTATGCTCAAAATCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 296165
QY 1021 CTTTACATTAATGAAGCAAGTGAAGCAAGTGAAGCAAGTGAAGCAAGTGAAGCAAGT 1080
Db 296164 CTTTATATCAATCAAGCAAGTGAAGCAAGTGAAGCAAGTGAAGCAAGTGAAGCAAGT 296105
QY 1081 GAAGTCGTTACTGATCAATTTCCATGACCCCTTACACCAAGCAAGCAAGCAAGCAAGT 1140
Db 296104 GAAGTCGTTACTGATCAATTTCCATGACCCCTTACACCAAGCAAGCAAGCAAGCAAGT 296045
QY 1141 AATGCTTTTATGCAACACCAAGTGGTATGATGACCAAGCAAGCAAGCAAGCAAGT 1200
Db 296044 AATGCTTTTATGCAACACCAAGTGGTATGATGACCAAGCAAGCAAGCAAGCAAGT 295985
QY 1201 GATGTAGATTAATGCAATTTTCTTATTAACCGCTTATTTAGAGCAATCAATTCGGTCCGCT 1260
Db 295984 GATGTAGATTAATGCAATTTTCTTATTAACCGCTTATTTAGAGCAATCAATTCGGTCCGCT 295925
QY 1261 TTTGGTTTTGAAGGTACACCAATTTCTATTTATGCTGCAAGAGAGA 1305
Db 295924 TTTGGTTTTGAAGGTACACCAATTTCTATTTATGCTGCAAGAGAGA 295880

RESULT 8
LOCUS AE017002
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
source

AE017002 304680 bp DNA linear BCT 16-MAY-2003
Bacillus cereus ATCC 14579 section 5 of 18 of the complete genome.
AE017002 AE016877
AE017002.1 GI:29894935
Bacillus cereus ATCC 14579
Bacillus cereus ATCC 14579
Bacteria: Firmicutes; Bacillales; Bacillaceae; Bacillus; Bacillus
cereus group.
1 (bases 1 to 304680)
Ivanova,N., Sorokin,A., Anderson,I., Galleron,N., Candelson,B.,
Kapatral,V., Bhattacharyya,A., Reznik,G., Mikhailova,N.,
Lapidus,A., Chu,L., Mazur,M., Goltzman,E., Larsen,N., D'Souza,M.,
Walunas,T., Grechkin,Y., Pusch,G., Haselkorn,R., Fonstein,M.,
Ehrlich,D.S.D., Overbeek,R. and Kyrpides,N.
Genome sequence of Bacillus cereus and comparative analysis with
Bacillus anthracis
Nature 423 (6935), 87-91 (2003)
22608415
12721630
2 (bases 1 to 304680)
Candelson,B., Gailoux,K., Ehrlich,D.S. and Sorokin,A.
The number of ribosomal RNA operons in Bacillus cereus
Unpublished
3 (bases 1 to 304680)
Ivanova,N., Sorokin,A., Anderson,I., Galleron,N., Candelson,B.,
Kapatral,V., Bhattacharyya,A., Reznik,G., Mikhailova,N.,
Lapidus,A., Chu,L., Mazur,M., Goltzman,E., Larsen,N., D'Souza,M.,
Walunas,T., Grechkin,Y., Pusch,G., Haselkorn,R., Fonstein,M.,
Ehrlich,D.S.D., Overbeek,R. and Kyrpides,N.
Direct Submission
Submitted (12-MAR-2003) INRA, Genetique Microbienne, Domaine de
Vilvert, Jouy en Josas 78352, France
Location/Qualifiers
1. 304680
/organism="Bacillus cereus ATCC 14579"
/mol_type="genomic DNA"
/strain="ATCC 14579"
/db_xref="ATCC:14579"
/db_xref="taxon:226900"
68. 427
/locus_tag="BC1239"
68. 427
/locus_tag="BC1239"
/codon_start=1
/transl_table=11
/product="hypothetical Membrane Spanning Protein"
/protein_id="AAP08224.1"
/db_xref="GI:29894936"
/translation="MLRRKLLYFLTPVLYAVLISMTKIELMALFLGVYVIFSNLRI
QESITLCLFSISIEFSIVSVILLNELFLRHSFTLKKFNIIQLAICITVIVFVL
KEIVGKTVEODNRKE"
551. 2041
/locus_tag="BC1240"
551. 2041
/locus_tag="BC1240"
/codon_start=1
/transl_table=11
/product="L-lectate permease"
/protein_id="AAP08225.1"
/db_xref="GI:29894937"
/translation="MLALIFIMIFICFLFKQTSRLRSLSIYAVSGVILLSPTFOL
GISCTVATNGMLCIVGVLFPGFIFLHNMNGYLDQVARELEVYTDORLQML
LMCGICPLIESVSGFGIPWAAPIFLSLGKPFQAVLLSFLGLASGWMATGTI
IGSOLINHPITLTGTTALLSDIPNFAYFVILSHVVGQNAVIEKRWGCVGFLLSL
GIYLSNAVSVELAGILSSITVTFGLIKLKGSGONLITEHAATERTGVSIKII
SPYIFLTVCLLSRLPALDPRSVAVLDKASYKLELYSPGLMGTCFTIIF
FRPSNIKQSLSQTIKQIPFAITTTTFIAISLWAGSGHSLAKTAGTGTCTFFV
FVAPFICAGFLTGSNAGSNANFINKQMQTAQNALPHQYVTTTLQNTASSVATACP

244138	CGTATTTATAGCGCTGGAGAAATGGTTAAATCATGAATTTAACATATATTGATACGGCGGA	244191
181	ATTGAAATTTGGTATGCACCACTTCCAAACACAAATATAGACGCGCAGCAGAAATTCGCCATA	240
244198	ATTTATTTTATTTGGAGACAGCGCGTTTTTGACACAAATTCGTCAACAGCTGAAGTAGCAAT	244257
241	GATGAAGCGGATGTTATTTATTTATTTATGCGTTAACGTCGTCGAAGGATTGACACAAAGCGAT	300
244258	GATGAAGCGGATGTTATCATTTTTTATGACAAACGTCGTCAACGCTGTAACGACGCTGAT	244317
301	GAATGGTCGCTCAATTTTATACAAATCTTAAAAACCGCTGCTATTTACGGTTTAAACAA	360
244318	GAAGAATTTTGCAAAATTTTATTCGTTTCAAATTAACCGGTTGCTACTGCAGTAAATAAA	244377
361	GTAGATATATGCGAAATCGGTACAGACGCTGTATGATTTCTATTATTAGGATTTTGGTGAA	420
244378	GTAGACAATTCGGAATTCGCGAGTGCATTTTATGACTTTTATGCATTTAGGATTTTGGCGAA	244437
421	CGGTATCCAATATACAGGCTCACATGGTTTATAGTCTTGGTCGACTTGTATTAGATGCACTTGT	480
244438	CCATTTCCCAATTTTCAGGTACACAGGATTAGGGTTAGTGCATTTACTAGACGACGCTGCA	244497
481	TCTCATTTTGGTGAAGAGGAAGACATCCTTATGATGAAGATACAAATTCGACTATGCCATT	540
244498	CAACATTTTCCAAAGATTTGAAGAGACGGATATGATGAAGACAAATTCGTTTCTCTTTA	244557
541	ATTGGACACCAACGTAAGTAAATCAAGTTTATGTAATGCTATTTTAGTGTGAAGATCGC	600
244558	ATTGGACGTCCAAACGTAAGGAAATCATCACTTGTAAATGCTCTTCTTGGTCAAGAGCGT	244617
601	GTTATCGTTTCTTAATGTTTGCAGGACACGAGACGCTATTTCATACAGGATATAGTTAT	660
244618	GTAAATGTTTATGTATGACGAGGAACAACCTCGTGATGCTGCTGATACACCATATAGTAA	244677
661	GATGCAACAGATTATGTTTTTAATTCGATACTGCTGGTATGCGTAAAAAGGAAAAAGTATAT	720
244678	GATGGAAGAAGATTATGTAATTTATGTATACAGCTGGTATGCGTAAAAAGGAAAAAGTGTAT	244737
721	GAATCAACTGTGAGAAATATTCAGTATTAAGAGCTTTTAAAGCGGATTTGAACGTTTCAATGTT	780
244738	GAARGTCAGAGAAAATATATAGTGTACTTTCGTGCGCTAAGAGCGGATTTGAACGTTTCTGACGTT	244797
781	GTTTATGTTGGTTATTGATGTCAGAACAAAGGCAATTTTGAACAAGATTAACCGTGTTCGACGA	840
244798	GTTTATGTCGTTTATGATGGTGAAGAGGAATTTATGAGCAAGATTAATAAATCCTCTGGA	244857
841	TATGCACATGAACAGGTTAAAGCAGTCTGTATGTCGTAAATTAATGGGATACTGTGGA	900
244858	TATGCTCATGATTTCAGCGGAGCTGTGTTTATCGTTGTAAATTAATGGGATGCAGTGAA	244917
901	AAAGATAGTAAACGATGAAGAAATTTTGAAGATGTAAGTACGTAAAGAAATTCCAATTTTFA	960
244918	AAAGACGAAAAAACAATGAAGACATTTTGAAGAAAAATATCCGTCTCATTTCCAAATCTTA	244977
961	GATTATGCACAAATTTGCTTTTGTCTGCTTAAAGACGACACAGATTTACGTACATTATTC	1020
244978	GAATACGCCCAATTTGTTTCTTATCTGCAAAAACGGAAGACGTACACAAACATTAATA	245037
1021	CCTTACATTTATGAAGCAAGTGAACCAATAAAAAAGCTGTTCCAAAGTTTCAACTTTAAAT	1080
245038	CCGGTTATTCGATCAGGTAAATGAAGTCAATAGTATCCGTATTTCCAAACGAACGTATTAAC	245097
1081	GAAGTCGTTTACTGATGCAATTTCCATGAACCCCTACACCAACACACAAAGGTAGAGCTTGT	1140
245098	GATGTAATTTATGATCGCGTAGCCGATGAATCCGAGGCCAACTCATTAATGTAAGCGTCTG	245157
1141	TAATGCTTTTATGCAACACAGTTGCTATAGACCACCGACATTTCTGTTATTTGTTAAT	1200
245158	AAATTTCTTATGCAACACAAGTTGCGGTAAACCCCGACATTTCTGTTGTTTGTAAAC	245217
1201	GATGTAGAAATTAATGCATTTTTTCTTTATAAACGCTATTTAGAGAAATCAAAATCCGTGCCGT	1260
245218	GATCTGTAATTTATTCACCTTCTCATATGAGCGTTTCTTAAGAAATCCGTTAGCTGAATCA	245277

OY 1261 TTTGGTTTTGAAGTACACCATTTCATATTATGCTCGAAGAAGA 1305
||||| || | || | || | || | || | || | || | || | || |
DB 245278 TTTGGCTTGAGAACGCCGATTTCACATTATGCTAGACAAGA 245322

```

RESULT 9
LOCUS AE017028
DEFINITION Bacillus anthracis str. Ames section 5 of 18 of the complete genome. BCT 30-APR-2003
ACCESSION AE017028
VERSION AE016879
KEYWORDS GI:30255149
SOURCE Bacillus anthracis str. Ames
ORGANISM Bacillus anthracis str. Ames
Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus; Bacillus cereus group.
REFERENCE 1 (bases 1 to 290117)
AUTHORS Read,T., Peterson,S., Tourasse,N., Baillie,L., Paulsen,I., Nelson,K., Tettelin,H., Fouts,D., Eisen,J., Gill,S., Holtzapple,E., Oksstad,O., Helgason,E., Rilstone,J., Wu,M., Kolonay,J., Beanan,M., Dodson,R., Brinkac,L., Gwinn,M., DeBoy,R., Madupu,R., Daugherty,S., Durkin,A., Haft,D., Nelson,W., Peterson,J., Pop.M., Khouri,H., Radune,D., Benton,J., Mahamoud,K., Jiang,L., Hance,I., Weidman,J., Berry,K., Plaut,R., Wolf,A., Watkins,K., Nierman,W., Hazen,A., Cline,R., Redmond,C., Thwaite,J., White,O., Salzberg,S., Thomas,B., Friedlander,A., Kochler,T., Hanna,P., Kolsto,A.-B. and Fraser,C.
TITLE The genome sequence of Bacillus anthracis Ames and comparison to closely related bacteria
JOURNAL Nature 423 (6935), 81-86 (2003)
MEDLINE 22608414
PUBMED 12721629
REFERENCE 2 (bases 1 to 290117)
AUTHORS Read,T., Peterson,S., Tourasse,N., Baillie,L., Paulsen,I., Nelson,K., Tettelin,H., Fouts,D., Eisen,J., Gill,S., Holtzapple,E., Oksstad,O., Helgason,E., Rilstone,J., Wu,M., Kolonay,J., Beanan,M., Dodson,R., Brinkac,L., Gwinn,M., DeBoy,R., Madupu,R., Daugherty,S., Durkin,A., Haft,D., Nelson,W., Peterson,J., Pop.M., Khouri,H., Radune,D., Benton,J., Mahamoud,K., Jiang,L., Hance,I., Weidman,J., Berry,K., Plaut,R., Wolf,A., Watkins,K., Nierman,W., Hazen,A., Cline,R., Redmond,C., Thwaite,J., White,O., Salzberg,S., Thomas,B., Friedlander,A., Kochler,T., Hanna,P., Kolsto,A.-B. and Fraser,C.
TITLE Direct Submission
JOURNAL Submitted (26-MAR-2003) The Institute for Genomic Research, 9712 Medical Center Dr, Rockville, MD 20850, USA
FEATURES source
1 . 290117
/organism=Bacillus anthracis str. Ames'
/mol_type='genomic DNA'
/strain='Ames'
/db_xref=taxon:198094'
331..1929
/locus_tag='BAI197'
331..1929
/locus_tag='BAI197'
/note=similar to GP:10717143, and GP:10717143; identified by sequence similarity; putative'
/codon_table=1
/transl_table=1
/product='oligopeptide ABC transporter, oligopeptide-binding protein'
/protein_id='AAP25159.1'
/db_xref='GI:30255150'
/translators='MMLGACSVKQDEPOANKGDSCSKAKOVILNTEVSEIPTMTDT LASADRSKYWNNTMEGLIRLKGDLPGVAKSHTKSDCKKYFKLRDAKNMGD LPTAKDYAWORAPNDTRAKSGPYIMLVOKNAIKRELSPQGLGVRAIDITVEF ELDPNPVFIDLTVLPFFPLNEKYMKEOGEKFLEANTLLTYNGPVNLNMKKHQDF LKRNPSZDOEKVLAEVFNWVYGTAALNVLSIAIDRVLGSAELVDLYKQGDSQF TINDTPVFRFLNRKPALKANIRKAITSAYDRVGIGETVLNNNSKGAYKGLVPDSF KGPQKOFRENGKSLKDVKWEAKMGKELGDKIEIEIFLNFOEESKRIGYE'
gene
CDS

```


Db	280434	AAATGAATTTGTTGGAGAGAGAGTTTCAATCGTAGAGATATATCCAGGTCTAAACGCCGAT	280493
QY	121	CGTATTTATCTTCAGGTGAGTGTAAACACATGATTTCAATATATTATGATACAGTGGT	180
Db	280494	CGTATTTATACCGTGGAGATGTTAAATCATGATTTAAATCATGATTTATGATACAGGTGGA	280553
QY	181	ATTGAATTTGTTGATGACCAATTCACAAACAAATTTAGAGCGCGAGCAGAAATTCGCCATA	240
Db	280554	ATTGATATTGGAGACGAGCCATTTTTCACACAAATTCGTCACAAAGCTGAAGTAGCAATT	280613
QY	241	GATGAACGGGTGTTATTTATTTTATGTTTATCGTGGTGAAGGATTGACACAAAGCGAT	300
Db	280614	GATGAACGAGATGTTATCATTTTATGACAAATGCTGCGGCTGAACACACGCGAT	280673
QY	301	GAATGTCGCTCAAAATTTTATACAAATCTAAACACCGTGTGATTTAGCGGTAAACAAA	360
Db	280674	GAAGAAGTTGCAAAATATTATATCGTTCAAAATTAACACCAAGTTGTACTTCGACGTAATAA	280733
QY	361	GTAGATAATATGGAATCCGTACAGAGCGTGTATGTTCTTATTCATTAGGATTTGGTGA	420
Db	280734	GTGGACAATCCGGAATGCCAGTGACATTTATGATTTCTATGCAATTAGGATTTGCCGAG	280793
QY	421	CGGTATCCAATATCAGGTGCATCGTTTAGGCTCTGGTGACTGTTAGATCGAGTGT	480
Db	280794	CCATTCCAATCTCAGGTACACACCGGATTAGGATTTAGGTGACTTTATAGTGAAGCTGCA	280853
QY	481	TCTCATTTTGGTGAAGAGGAAGATGCTTTATGATGAAGATACAAATTCGACATTCAT	540
Db	280854	CAGCATTTTCCAAAGATTTGAAGAAGACGGATATGATGAAGATACAAATTCGTTCTCTTTA	280913
QY	541	ATTGACGACCAACAGTAGGTAAATCAAGTTTATGTAATGCTATTTTAGTGCAAGATCGC	600
Db	280914	ATTGCGCGTCCAAATGTAGGAAATCATCTTGTAAATGCTCTCTTGGTCAAGAGCGT	280973
QY	601	GTTATCGTTCTTAATGTTGAGGACACAGAGAGCGTATTTGATACAGATATAGTAT	660
Db	280974	GTAATGTTAGTATGACGAGGAACCACTGCTGATGCTGATACACCATATAGTAAA	281033
QY	661	GATGACAAGATATGTTTTAATCGATCTGCTGCTGATGCTGTAAGGAAAAAGATAT	720
Db	281034	GATGAAAAGATATGTAATTTATCGATACAGCTGCTGCTGTAAGGAAAAAGGATGAT	281093
QY	721	GAATCAACTGAGAATATTCAGTATTAAGAGCTTTTAAAGCGATTGAACGTTTCAATGTT	780
Db	281094	GAAGTACAGAAAAGTATAGTACTTCTGCTGACATGAAGCGATTTGACGCTTCACT	281153
QY	781	GTTTTAGTGGTATTGATGACAGAACAGGCAATTTATGACACAGATAAACGTTTCCAGGA	840
Db	281154	GTTTTAGTGGTATTGACGCGTGAAGAGGAATTTATGAAACAGATAAAAAATTCCTGGA	281213
QY	841	TATGACATGAACAGGTAAAGCAGTCTGATTTGCTGTAATTAATGCTGATCTGCGAA	900
Db	281214	TACGCTCATGATTCAGGACGAGCTGCTGATCGTTGTTAAACAAATGGACGACGTGAA	281273
QY	901	AAAGATAGTAAACGATGAAGAATTTGAAGATGAAGTACGTTAAAGAATTCCAATTTTA	960
Db	281274	AAAGATGAAGAACAATGAAGCAATTTGAAGAACAATTCGCTGCTATTTCCAAATTTTA	281333
QY	961	GATTATGCAAAATTCGTTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1020
Db	281334	GAGTACCGCAGCGATTTGATTTCTTATCTGCAAAACGAGAAAACGCTACACAAACGTTA	281393
QY	1021	CCTTACATTAATGAACGATGAACCAATTAACCAATTAACCAATTAACCAATTAACCAAT	1080
Db	281394	CGGTTATCGATGAGTAAATGAAGCCATAGTATCCGTTATTAACCAAGATTTAAT	281453
QY	1081	GAGTCTGTTACTGATGCAATTTTCATGACCCCTACACCAACAGCAAAAGGTAGCGTTG	1140
Db	281454	GATGTAATATGATGATGCGGTAGCGATGAATTCACACCCCAACGATAGTGTGCTG	281513
QY	1141	AATGTCGTTTTATGCAACAAAGTTGCTATAGAACCACCCGACATTTGTTGTTATTTAAT	1200
Db	281514	AAGATTTTCTATGCAACACAAGTTGCGGTAAACACCAACCAATTTGTTGTTGTTAAT	281573
QY	1201	GATGTAGATTAATGCATTTTCTTATATAACCGCTATTATAGAGATCAAAATCGTCCGCT	1260
Db	281574	GATCCAGAACTATTGCATTTCTCATATGAGCGCTTCTTAAAAAATCGTTACGTGAATCA	281633
QY	1261	TTTGGTTTTGAAAGGTACACCAATTCATATATAGCTCGAAAGAGA	1305
Db	281634	TTCCGCTTTGTAGGAACGCCGATTCACATTTATCGTAGCAAGA	281678
RESULT 10			
AL591981/c			
LOCUS			
DEFINITION			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
COMMENT			
FEATURES			
source			
gene			
CDS			
gene			
CDS			


```
Db 259438 ATGGCAAAACCAAGTTGTAGCGATTGTGCGAGCTCCAAACAGTTGGCAAAATCGACTATTTT 259379
Qy 61 AATAGAAATAGTTGGAGAACGTTTTCGATTGTTGAAGATACGCCAGCTGAACACAGAT 120
Db 259378 AACAGAAATCGTTGGTGAACGTTTCCATAGTGAAGATGTTCCGGTGTGACACGTGAC 259319
Qy 121 CGTATTATTCTTCAGGTGAGTGTAAACACATGATTTCATATTATTATGATACAGGTGT 180
Db 259318 CGCATATATATTCAGCCGAATGGCTTGGAAAGAATTTAACATATTATTATGACAGGTGT 259259
Qy 181 ATTGAATTTGGTGCACCAATTTCCAAACACAAATTAGAGCGCAGGAGAAATCGCCATA 240
Db 259258 ATTGATCTTTCCGATGAACCAATTTTAGAGCAAAATTCGCGCACAGCGAAATCGCAAT 259199
Qy 241 GATGAAGCGATGTTATTATTATTATGTTTAACTGCGTGAAGGATGACAAACCGAT 300
Db 259198 GATGAAGCAGAGCTAATTTATTATTATACCAGTGTGCTGAAGGATTCACGATCGAGAC 259139
Qy 301 GAATGTCGCTCAAAATTTATACAAATCTAAAAACCGTGTGCTATTAGCGTTAAACAAA 360
Db 259138 GAACAAGTAGCAAAATTTCTTACCGGTCTAATAAACCAATTTGTTTAGCGATTAAATAA 259079
Qy 361 GTAGATAATATGGAATCGGTACAGACGTGTATGATTCTATTCAATTAGGATTGGTGAA 420
Db 259078 GTAGATAACCCAGAAATCGAGATCAAAATTTATGACTTTTATTCTCTGGTGTGGAG 259019
Qy 421 CGGTATCCAAATACAGGCTACATGGTTTAGGTCTGTGTCACGTGTAGATCGATGTT 480
Db 259018 CGGTATCCAAATTTCTGTTACATGAGCTAGGCTGTGGCAGACATGCTTGATGCTGCC 258959
Qy 481 TCTCATTTGGTGAAGGAAGAGATCCTTTATGATGAAGATACAAATTCGACTTCATTT 540
Db 258958 GCTCATATTTCCAAAGAGAAGAGAGAGAAATATCCAGATGACACAGTGAATTCAGTTTA 258899
Qy 541 ATTGGACGACCAACAGTAACTAGTAAATCAAGTTTAGTAAATGCTATTTTAGTGAAGATCG 600
Db 258898 ATCGGTGCGACCAATGTTGGTAAATCTTATCTTAATCACTGCTTGGAGAGATCGT 258839
Qy 601 GTTATCGTTCTTAATGTTGCGAGGACGACGAGACGCTATTATACAGATGATGTTAT 660
Db 258838 GTTATGTTTCAGATTATTCGGGTACACCTGCTGATGCAATGATACAACTTATACCTTC 258779
Qy 661 GATGGACAAGATTATGTTTTTAATCGATCTGCTGTTATGCGTAAAAAGGAAAGATATAT 720
Db 258778 GATGGCCAGATATTATGTCATGATTGATACAGCCGAATGAGAAACGTTGGAAAGTGTAT 258719
Qy 721 GAATCAACTGAGAAATATTCAGTATTAAAGAGCTTTTAAAGCGATTGCAAGCTTCAATGTT 780
Db 258718 GAAACACAGAGAAATATAGTGTGTTTACGTGCAATGAGACCAATTCAGCTCGGACGTT 258659
Qy 781 GTTTTAGTGGTTATTGTCAGACAGAACAGGCAATTTTGAACAGATAAACGTTGTCAGGA 840
Db 258658 GTTCTGTGGTTATCAAGCGAGAGAAAGGTTATTCGTGAGCAGATTAACGCAATTCGTGA 258599
Qy 841 TATGCATATGAACAGGTAAAGCAGTCTGATTGCTGTAATTAATAGGATAGTGTGGA 900
Db 258598 TATGCCCATGATGCCGACGTGCAATTTATTATTGTTGTAACAAATGGGATCAATTAAC 258539
Qy 901 AAAGATAGTAAACGATGAGAAATTTGAAGATGAGTACGTAAAGCAATTCCAATTTTA 960
Db 258538 AAGATGAAANACAAATTAATGATGACAGAGATATTTCGTGAGCAATTCCAATTCCTTA 258479
Qy 961 GATTATGCACAAATTTGTTGCTGCTGTGCTTAAAGACGCAAGATTTAGCTACATTTTC 1020
Db 258478 AGCTATGCACCAATTTGTTGCTATCTGCTTAAACAAACAAACGCTTAAATATCTATTC 258419
Qy 1021 CCTTACATTAATGAGCAAGTAAACCACTTAAACAAACGCTTCAAAAGTTCAACTTTAAT 1080
Db 258418 CCACCTATTAAACCAAGTAAAGGATACCACTTCAATACGTGATCAATCAAGTATGCTGAT 258359
Qy 1081 GAAGTGTGTTACTGATGCAATTTCCATGAACCCCTACCAACAGACAAAGGTAGAGGTTTG 1140
Db 258358 GATGTTATTAGTGTGCTGTGCAATGAACCCATCACCTATGATTAAGGTAAACGACTG 258299
```

1141 AATGCTCTTTTATCAACACACAGTTGCTATAGAACACACGACATTTGTTGTTATTTTAAAT 1300
1141 AATGCTCTTTTATCAACACACAGTTGCTATAGAACACACGACATTTGTTGTTATTTTAAAT 1300
258298 AAGATTTCTTATCAACACACAGTTGGGGTAAACACACCAAGCTTTGTTGTTATTTTAAAT 258239
1201 GATG7AGAATTAATGCAATTTTCTTATATAAACCCTATTATTAGAGAATCAAAATCCGTCGCCCT 1360
258238 GATCCAGACATCATGCAATTTTTCGTATGAACGTTTCTTTAGAAAACCGGATAGAGAAGCA 258179
1261 TTTGGTTTTCAGGTCACACAAATTCATATTATGATCTCGAAAGAG 1304
258178 TTTCCATTTCAGGTCACGCAATTCGAGTAAATGCTGTAAGCG 258135

RESULT 11
AX641670/c
LOCUS AX641670
DEFINITION Sequence 2860 from Patent WO0101118.
ACCESSION AX641670
VERSION AX641670.1 GI:28474431
KEYWORDS
SOURCE Listeria monocytogenes
ORGANISM Listeria monocytogenes
REFERENCE 1
AUTHORS Buchrieser, C., Frangeul, L., Couve, E., Rusnok, C., Falh, H., Dehoux, P., Dussurget, O., Chetouani, F., Nedjari, H., Glaser, P., Kunst, F., Cossart, P., Daniels, J., Goebel, W., Kref, J., Kuhn, M., Ng, E., Vasquez-Boland, J., Dominguez-Bernal, G., Garrido-Garcia, P., Tierrez-Martinez, A., Amend, A., Chakraborty, T., Domann, E., Hahn, T., Berche, P., Charbit, A., Durant, L., Perez-Diaz, J. C., Baquero, F., Garcia del Portillo, F., Gomez-Lopez, N., Madueno, E., de Pablo, B., Wehland, J., Kierst, U., Entian, K. D., Hauf, J., Rose, M., and Voss, H.
TITLE Listeria monocytogenes genome. Polypeptides and uses
JOURNAL Patent: WO 0101118-A 2860 11-APR-2001;
FEATURES
Location/Qualifiers
source
1. 349980
/organism="Listeria monocytogenes"
/mol_type="genomic DNA"
/db_xref="taxon:1639"
/note="Seq 1 for big; original length: 2944528-split
into 10 <23>seq 0001 : From 0.000.001 to 0.349.980
length: 349.980 <23>seq 2855 : From 0.300.001 to
0.649.980 length: 349.980 <23>seq 2856 : From 0.600.001
to 0.949.980 length: 349.980 <23>seq 2857 : From
0.900.001 to 1.249.980 length: 349.980 <23>seq 2858 :
From 1.200.001 to 1.549.980 length: 349.980 <23>seq 2859
: From 1.500.001 to 1.849.980 length: 349.980 <23>seq
2860 : From 1.800.001 to 2.149.980 length: 349.980
<23>seq 2861 : From 2.100.001 to 2.449.980 length:
349.980 <23>seq 2862 : From 2.400.001 to 2.749.980
length: 349.980 <23>seq 2863 : From 2.700.001 to
2.944.528 length: 244.528"
BASE COUNT 99761 a 75437 c 61177 g 113605 t
ORIGIN

Query Match 52.4%; Score 683.2; DB 6; Length 349980;
Best Local Similarity 70.2%; Pred. No. 2.9e+109;
Matches 916; Conservative 0; Mismatches 388; Indels 0; Gaps 0;
1 ATGACTAAACCTATAGTACGTATTGTTAGGTAGGCTTAATGTAGTAAATTCACAAATTTT 60
212438 ATGGCAAAACCAAGTTGTAGCGATTGCGAGCTCCAAACGTTGCAAAATCGACTATTTT 212379
61 AATAGAAATAGTTGGAGAACGTTTTCGATTGTTGAAGATACGCCAGCTGAACACAGAT 120
212378 AACAGAAATCGTTGGTGAACGTTTTCATAGTGAAGATTTCCCGGTGTGACACGTTGAT 212319
121 CGTATTATTCTTCAGGTGAGTGTGTTAACACATGATTTCATATTATTGATACAGGTGT 180
212318 CCATATATAATTCACCGGAATGCTTGGAAAGAATTTTAAACATTTATGATACAGGTGT 212259

QY	121	CGTATTATCTTCAGGTGAGTGGTTAACACATGATTTCAATATTATTGATACAGGTGGT	180
Db	276690	CGCATATATATTCAGCGGAATGGCTTGGAAAGAAATTTAAACATTTATTGATACAGGTGGT	276631
QY	181	ATTGAATTTGGTGATGACACATTTCCAAACACAAAATTTAGAGCCGAGCGAAATGCCCAT	240
Db	276630	ATTGATCTTTCCGACGAACCATTTCTTAGAGCAAAATTCGGCCACAAAGCGGAAATGCCAAT	276571
QY	241	GATCAAGCGGATGTTATTTATTTATGTTTAAAGTCGTTGAAGGATTCGACAAAGCGAT	300
Db	276570	GATCAAGCAGACGCTATTTATTTATACCAATGGTGGTGAAGGGTTACCGATGCAGAC	276511
QY	301	GAATGTCCTGCTCAATTTTATACAAATCTAAAAACCGGTCGTATTAGCGGTTAAACAA	360
Db	276510	GAACAGTGAACAAAATCTTTACCGGCTCTAATAACCAATTTGTTTATAGCGATTAATAA	276451
QY	361	GTAGATAATATGGAATTCGGTCAGACGCTGATGTTCTATTCTATTAGGATTTGGTGAA	420
Db	276450	GTAGATAACCCAGAAATGCGTGATCAGATTTTATCTCTTTGGTGGTGGT 276391	
QY	421	CCGTATCCAAATATCAGGTCACATGTTTAGCTCTGTGACTTCTTAGATGCAATGTT	480
Db	276390	CCGTATCCAAATTTCTGTTCTCATGACTAGGCTTGGTGATATGCTTGATGCTGCTGGT	276331
QY	481	TCTCATTTTGGTGAAGGAGGAAGATCCTTATGATGAAGATACAAATTCGACTATCCAT	540
Db	276330	GCTCATTTTCCAAAGAAAGAGGAAGAAATACCCAGATGAACAGTGAATTTAGTTTG	276271
QY	541	ATTGCGAGCACAACCGTAGTAAATCAAGTTTAGTAAATGCTATTTTAGGTGAAGATCGC	600
Db	276270	ATTGCGTGGGCAAAATGTTGTAATCATCTATCTAAGCCACTTCTTGGCGAGACCGC	276211
QY	601	GTTATGCTTTCTAATGTCGAGGCAACGAGAGAGCTATTGATGACAGATGATGATAT	660
Db	276210	GTCATTTGTTCTGACATTCGGGAACACTCTCTGATGCAATTTGATACAACTTATACGTT	276151
QY	661	GATCGACAAGATTATGTTTAAATCGATCTGCTGATGCTGCTGCTGCTGCTGCTGCTG	720
Db	276150	GATGTCAGGATTTATGTCATGATTGATACAGCGGAATGAGAAACGTTGGCAAGTGTAT	276091
QY	721	GAATCAACTCAGAAATATTCAGTATTAAAGAGCTTTAAAGCGATTGAACGTTCAATGTT	780
Db	276090	GAAGCACACAGAAATATAGTGTTTACGTGCAATGAGAGCAATTTGAACGCTCCGATGTT	276031
QY	781	GTTTATGTTGTTATGATGACAGAACAGCATTTATTGAACAGATTAACAGTGTGTCAGGA	840
Db	276030	GTTCTTGTTGTTATCAACGACAGAAAGGTATTTCGTGAGCAAGATTAAGCGATTGCTGA	275971
QY	841	TATGCACATGAACAAGGTAAAGCAGTCGTGATTGTCGTAAATAAATGGGATCTGTGAA	900
Db	275970	TATGCGCATGATCCGCGCGCTATCATTTATTGATGAACAATGGGATGCAATTAAC	275911
QY	901	AAAGATAGTAAACGATGAAGAAATTTGAAGATGAAGTACGTAAAGAAATCCCAATTTTA	960
Db	275910	AAAGTGAACAAACAAATTTATGATGGACAGAAATATTCGGGAGCAGTTCCCAATTTCTA	275851
QY	961	GATTATGCAAAATGCTTTTGTGCTGCTGAAGAACGCAAGATTAACGATTAATTC	1020
Db	275850	AGCTATGACCAATTTGTTTCTGATCTGCTGAACAAACAAACGCTTAACAACTTATTC	275791
QY	1021	CCTTACATTAATCAAGCAAGTGAACACCATAAAAACGTTTCAAGTTCCACTTTAAAT	1080
Db	275790	CCGCTCATTAAACCAAGTGAAGCAATATCATTTACCGGTACATCTAGTATGCTAAAT	275731
QY	1081	GAAGTCGTTACTGCAATTTCCATGAACCTTACACCAAGCAAGATTAACGTTG	1140
Db	275730	GATGTTATTAGTATGCGGTTGCAATGAATCAATCAACCAATGGATTAAGGTTAAACGACT	275671
QY	1141	ANTGCTCTTTATGCAACACAAATTTGCTATAGAACCCGACGATTTGTTGTTATTTGTTAA	1200
Db	275670	AAATATTTCTATACACACAAAGTGGCTGTTAAACCCGCAACATTTGTTGTTATTTGTTAA	275611
QY	1201	GATGTAGAAATTAATGCAATTTTCTTATAAACGCTATTTAGAGAAATCAAAATCCGCGCCT	1260
Db	275610	GATCAGAACTAATGCAATTTTCTTATAAACGCTATTTAGAGAAATCAAAATCCGCGCCT	275551
QY	1261	TTTGGTTTGAAGGTACACCAATTTATATATATATATATATATATATATATATATATAT	1304
Db	275550	TTTCCGTTTGTATGGTACGCCAATTCGAGTAATTTGCTGCTGAAGCG 275507	
RESULT 13			
AX413016	Sequence	319630 bp	DNA linear PAT 03-SEP-2002
LOCUS	AX413016	Sequence 7 from Patent WO0228891.	
DEFINITION	AX413016	Accession	
VERSION	AX413016.1	GI:21445474	
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
FEATURES			
source			
BASE COUNT	105207 a	55428 c	66726 g 92263 t
ORIGIN			
Query Match		51.94%	Score 676.8; DB 6; Length 319630;
Best Local Similarity		69.94%;	Pred. No. 3.8e-108;
Matches	912;	Conservative	0; Mismatches 392; Indels 0; Gaps 0;
QY	1	ATGACTAAACCTATAGTAGCTATTGTAGTAGGCGCTAAATGTAGTAGTAAATCTACAATTTT	60
Db	50618	ATGGCAAAACCCAGTTGTAGCGATTGTGCGAGCTGCCAACGTTGGCAATTCGACTATTTT	50677
QY	61	AATAGATAGTTGGAGAACGTTTCCGATTGTTGAGATACGCCAGCGTACACAGAGAT	120
Db	50678	AACAGATCGTTGGTGAACGTTTCCATAGTGGAGAGATGTTCCCGGCTGACACGTGAC	50737
QY	121	CGTATTTATTTCCAGGTGAGTGTTAAACACATGATTTCAATATTATTGATACAGGTGGT	180
Db	50738	CCCATATATAATTCAGCGGAATGCTTTGGAAAGAAATTTAAACATTTATTCACAGTGGT	50797
QY	181	ATTGAAATTTGCTGATGCACCATTCACAAACACAAATTAGAGCGCAGCAAGATCCGATA	240
Db	50798	ATTGATCTTTCCGACGAACCATTCCTTAGACCAATTTCCGCGACACAGCGGAAATCCCAAT	50857
QY	241	GATGAAGCGGATGTTATTTTATGTTTAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	300
Db	50858	GATGAAGCAGACGTAATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTAT	50917
QY	301	GAATGTTGCTGCTCAATTTTATACAAATCTAAAAAACCGGTCGTATTAGCGGTTAAACAA	360
Db	50918	GAACAAGTAGCAAAATTTCTTACCGGCTCTAATAAACCATAATTTGTTAGCGATTAATAA	50977
QY	361	GTAGATAATATGGAATTCGGTACAGACGCTGATGTTCTATTCTATTAGGATTTGGTGAA	420
Db	50978	GTAGATAACCCAGAAATGCGGTGATCAGATTTATGACTTTTATTTCTCTTTGGTGGTGG	51037
QY	421	CCGTATCCAAATATCAGGTCACATGTTTAGGCTTTGGTGGTCTTTGGTGGTCTTTAGATG	480
Db	51038	CCGTATCCAAATTTCTGCTCTCATGGACTAGGCTTGGTGATATGCTGCTGCTGCTGCTGCT	51097
QY	481	TCTCATTTTGGTGAAGGAGGAAGATCCTTATGATGAAGATACAAATTCGACTATCCCAT	540
Db	51098	GCTCATTTTCCAAAGAGAGAGAGGAAGATACCCAGATGAACACTGTAATTTAGTTAGT	51157
QY	541	ATTGACGACCAACGCTAGGTAGTAAATCAAGTTTATAGTAAATGCTATTTTAGTGAAGATCGC	600

Db 51158 ATTGTCGCCCAATGTTGGTAAATCATCTATTCTAAACGCACCTCTTGGCGAAGACCGC 51217
Qy 601 GTTATGCTGTTCTTAATGTTGCGAGGACACGAGAGACGCTATTGATACAGAGTATAGTTAT 660
Db 51218 GTCAATGTTCTGACATTCGCGGAACTACTCGTGATGCAATGATGATCACTTATACGTTT 51277
Qy 661 GATGACACAGATATATGTTTAAATCGATACGCTGCTGATGCGTGAATGAAAGGAAAGTATAT 720
Db 51278 GATGCTGAGGATATGTCATGATGATACAGCGGAAATGAGAAACGCTGGAAAGGTAT 51337
Qy 721 GAATCACTGAGAAATATTCAGTATTAAAGAGCTTTAAAGCGATTGAAGCTTCAATGTT 780
Db 51338 GAAACACAGAGAAATATAGTGTGTTTACGTGCAATGAGAGCAATTAACCGCTCCGATGTT 51397
Qy 781 GTTTTAGTGGTTATGATGACAGAAAGGATATGATGAGCAATTAACCGCTCCGATGTT 840
Db 51398 GTTCTGTTGTTATCAGCGCAAGAAAGGATGTTGATGAGCAATTAACCGCTCCGATGTT 51457
Qy 841 TATGACATGACAAAGGTAAGCAGTGTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTG 900
Db 51458 TATGCGCATGATGCGGACGCGCTATCATTTATGATGAGCAAAATGGATGCAATTAAC 51517
Qy 901 AAAGATAGTAAACGATGAAGAAATTTGAAGATGAAGTACGTAAGCAATTCGAATTTTA 960
Db 51518 AAAGATGAAGAAACAAATTAATGATGACAGAGAGATATTCGGGAGCAGTTCGAATTTTA 51577
Qy 961 GATTATGACAAATGCTTTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1020
Db 51578 AGCTATGACCAATGTTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 51637
Qy 1021 CTTTACATTAATGAAGCAAGTGAAGCAATTTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 1080
Db 51638 CCGCTCAATTAACCAAGTGAAGCAATTTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 51697
Qy 1081 GAAGTCGTTACTGATGAATTTCCATCAACCTACACCAAGCAAGTGAAGTGAAGTGAAGTGA 1140
Db 51698 GATGTTATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 51757
Qy 1141 AATGCTTTTATGACACAAAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1200
Db 51758 AAAATATATCTATACAAACAAAGTGGCTGTTTAAACCGCAACATTTGTTGCTGCTGCTGCT 51817
Qy 1201 GATGATGAATTAATGCAATTTTCTTATTAACGCTATTTAGAGATCAAAATCCGCTCCGCT 1260
Db 51818 GATCCAGAACTAATGCAATTTCTTATGAAGCTTTCTTAGAAACCGGATTAGAAGCA 51877
Qy 1261 TTGCTTTTGAAGCTACACCAATTCATATATAGCTCGAAGAG 1304
Db 51878 TTTCGTTTGTATGCTAGCCCAATTCGAGTAAATGCTCGTAAGCG 51921

RESULT 14
AX417046/c 349980 bp DNA linear PAT 02-SEP-2002
LOCUS Sequence 4037 from Patent WO0228891.
DEFINITION AX417046
ACCESSION AX417046
VERSION AX417046.1 GI:21449656
KEYWORDS
SOURCE Listeria innocua
ORGANISM Listeria innocua
Bacteria; Firmicutes; Bacillales; Listeriaceae; Listeria.
REFERENCE 1
Kunst, F. and Glaser, P.
AUTHORS Listeria innocua, genome and applications
TITLE Patent: WO 0228891-A 4037 11-APR-2002;
JOURNAL INSTITUT PASTEUR (FR); CENTRE NATIONAL DE LA RECHERCHE
SCIENTIFIQUE (CNRS) (FR)
FEATURES Location/Qualifiers
source 1..349980
/organism="Listeria innocua"
/mol_type="genomic DNA"
/db_xref="taxon:1642"

/note="seq 2058, original length: 3,011,208 replaced
by-seq 2058: 0,000,001 to 0,349,980-seq 4032: 0,300,001 to
0,649,980-seq 4033: 0,600,001 to 0,949,980-seq 4034:
0,900,001 to 1,249,980-seq 4035: 1,200,001 to
1,549,980-seq 4036: 1,500,001 to 1,849,980-seq 4037:
1,800,001 to 2,149,980-seq 4038: 2,100,001 to
2,449,980-seq 4039: 2,400,001 to 2,749,980-seq 4040:
2,700,001 to 3,049,980-seq 4041: 3,000,001 to 3,011,208"
BASE COUNT 101055 a 72969 c 50688 g 115268 t
ORIGIN
Query Match 51.9%; Score 676.8; DB 6; Length 349980;
Best Local Similarity 69.9%; Pred. No. 3,7e-108;
Matches 912; Conservative 0; Mismatches 392; Indels 0; Gaps 0;
Qy 1 ATGACTAAACCTATATAGTACCTATGCTAGGTAGGCTTAATGTAGGTAAATCTACAATTTT 60
Db 266810 ATGGCAAAACAGTTGTAGCGATTGTCGGACGTCCTAAACGCTGCAAAATCGACTATTTT 266751
Qy 61 AATAGAAATAGTTGGAGAACGCTGTTTCGATTTGTTCAAGATACGCCAGGTGTAAACACGAGAT 120
Db 266750 AACAGAAATCGTTGGTGAACGCTGTTTCCATAGTGAAGATGCTCCCGGTGTGACACGCTGAC 266691
Qy 121 CGTATTTATTTCTTCAAGTGTAGTGTAAACATCATTTCAATATTATTTATGATACAGGTGT 180
Db 266690 CGCATATATTAATTCAGCGGAATGGCTTGGAAAGAAATTTAAACATTTATGATACAGGTGT 266631
Qy 181 ATTGAAATTTGGTGTGACACCATTTCCAAACACAAATTTAGAGCGGACGAGAAATTCGCCATA 240
Db 266630 ATTGATCTTTCCGAGGAACCATTTCTAGAGCAAAATTCGCCACAGCGGAATTCGCAAT 266571
Qy 241 GATGAAGCGGATGTTATTTATTTATGTTTAACTGCTGAAGGATTTGACACAAACGAT 300
Db 266570 GATGAAGCAGAGCTAATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTAT 360
Qy 301 GAATGCTGCTCAAAATTTTATACAAATCTAAAAACCGGTGCTATTATGCGGTAAACAAA 360
Db 266510 GAACAAATGAGCAAAATTTTACCGGTCTAATTAACCAATTTGTTTACGATTAATAAA 266451
Qy 361 GTAGATAATATGGAATTCGTACAGAGCTGTATGATTTCTATTCTATTAGGATTTGCTGAA 420
Db 266450 GTAGATAACCCAGAAATTCGTGATCAGATTTATGACTTTTATCTCTCTGCTGCTGAG 266391
Qy 421 CGTATTCAAATATCAGGCTCACATGTTTACGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 480
Db 266390 CGTATTCAAATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 266331
Qy 481 TCTCATTTTGGTGAAGAGGAAAGATGCTTTATGATGAAGATACAAATTCGACTATCCATTT 540
Db 266330 GCTCATTTTCCAAAAGAGAGGAAAGATACCCAGATGAAGAAACGTTGAAATTTAGTTTG 266271
Qy 541 ATTGGACGACCAACGCTAGGTAAATAGTAAATGCTATTAGTAAATGCTATTAGCTGAAGATCG 600
Db 266270 ATTGGTCGCCAAATTTGTTGTTAAATCATCTATTCTAAACGCACTTTCTTGGCAAGACCG 266211
Qy 601 GTTATGCTTCTTAATGTTGCGAGGACCAACAGAGACGCTATTGATACAGATAGTTAT 560
Db 266210 GTCATTTGTTGACATTTGCGGAACTACTCGGTGATGCAATTTGATACAACTTATAGTTT 266151
Qy 661 GATGACAAAGATTTATTTTAAATCGATGCTGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCT 720
Db 266150 GATGCTCAGGATTTATGTCATGATGATACAGCGGAATGAGAAACGCTGGAAAGTAT 266091
Qy 721 GAATCACTGAGAAATATTCAGTATTAAAGAGCTTTAAACGCTTTAAACGCTTTCAATGTT 780
Db 266090 GAAACACAGAGAAATATAGTGTGTTTACGTCGAATGAGAGCAATTTGACGCTCCGATTT 266031
Qy 781 GTTTTAGTGGTTATGATGCAAGAAAGGATTTATTTGAACAGATAAACGCTGCTCAGGA 840
Db 266030 GTTCTTGTGTTATCAACGCGAGAGAGGATTTCTGAGCAAGATTAAGCGGATGCTGGA 265971
Qy 841 TATGACATGAACAGGTTAAAGCAGTCTGATTTGCTGTAATTAATTTGGATCTCTGGA 900
Db 265971

Db	265970	TATGGCGATGATCGCGAGCGGCTATCATTTATTTGTAGTGAACAAATGGGATGCAATTAAC	265911
Qy	901	AAAGATAGTAAACGATGAAGAAATTTGAAGATGAAGTACGTAAAGAAATTCCAATTTTAA	960
Db	265910	AAAGATGAACAAATTAATGTATGGACAGAAAGATATTCGGGAGCAGTTCCAAATTCCTTA	265851
Qy	961	GATTATGCACAAATGCTTTTGTGCTGCTAAAGAACGACCAAGATTAAGTACATATTTC	1020
Db	265850	ACGTCACCAATTTGTTTGTGCTGCTAAACCAACCAACGCTTAACCAACTTATTC	265791
Qy	1021	CCTTACATTAATGAAGCAAGTGAACCACTAAACCAAGCTGTTCAACAGTTCAACTTTAAAT	1080
Db	265790	CCGCTCATTAACCAAGTAAGGATCAATTCATTCATTCACGCTACAATCTATAGTATGCTAAAT	265731
Qy	1081	GAAGTCGTTACTGATGCAATTTCCATGAACCCCTACACCAACAGCAAAAGGTAGAGCTTTG	1140
Db	265730	GATGTTATTAGTATGCGGTTGCAATGAATCCATCACCATGGAATGAAGGTAAACGACTT	265671
Qy	1141	AATGCTTTTTATGCAACACAAAGTTGCTATAGAACCACCGACATTTCTGCTATTTCTTAAT	1200
Db	265670	AAATATTTCTATACACACAAAGTGGCTGTTAAACCGCCCAACATTTGTTGTTATTTCTTAAT	265611
Qy	1201	GATGTAGTAATTAATGTCATTTTCTTATTAACCGCTATTATTAGAGNATCAATCCGTCGCCCT	1260
Db	265610	GATCCAGAACATTAATGTCATTTCTTATGAACGTTCTTATGAACCGGATTTAGAGAAGCA	265551
Qy	1261	TTTGGTTTTGAAGGTACACCAATTCATATTATAGCTCGAAAGAG	1304
Db	265550	TTTCCGTTTGTGATGGTACGCCAATTCGAGTAATTTGCTCGTAAGCG	265507
RESULT 15			
LOCUS	AX144037	960 bp	DNA
DEFINITION	Sequence 2759 from Patent WO0134809.	linear	PAT 31-MAY-2001
ACCESSION	AX144037		
VERSION	AX144037.1	GI:14282890	
KEYWORDS	synthetic construct		
SOURCE	synthetic construct		
ORGANISM	artificial sequences.		
REFERENCE	1		
AUTHORS	Kimmerly, W. J.		
TITLE	Staphylococcus epidermidis nucleic acids and proteins		
JOURNAL	Patent: WO 0134809-A 2759 17-MAY-2001;		
GLAXO GROUP LIMITED (GB)	Location/Qualifiers		
FEATURES	1..960		
source	/organism="synthetic construct"		

Search completed: September 12, 2003, 19:27:32
Job time : 5331.98 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 12, 2003, 17:17:59 ; Search time 1041.84 Seconds
(without alignments)
9028.139 Million cell updates/sec

Title: US-09-815-242-1463

Perfect score: 387
Sequence: 1 gatctctctcttcacca.....ttcacctgaagaataaatac 387

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*

1: em_estba:*

2: em_esthum:*

3: em_estlin:*

4: em_estnu:*

5: em_estov:*

6: em_estpl:*

7: em_estro:*

8: em_htc:*

9: gb_estl:*

10: gb_est2:*

11: gb_htc:*

12: gb_est3:*

13: gb_est4:*

14: gb_est5:*

15: em_estom:*

16: em_estom:*

17: em_gss_hum:*

18: em_gss_inv:*

19: em_gss_pln:*

20: em_gss_vrt:*

21: em_gss_fun:*

22: em_gss_mam:*

23: em_gss_mus:*

24: em_gss_pro:*

25: em_gss_rod:*

26: em_gss_phg:*

27: em_gss_vrl:*

28: gb_gss1:*

29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	159.4	41.2	2662	28	BH770770 LLMGtag51
C 2	93.4	24.1	552	9	A1390536 mu83B02.y
C 3	93	24.0	431	9	AA215138 mu83B02.f
C 4	51	13.2	525	12	BM274963 PREStoaar

C 5	46.6	12.0	1101	29	CNS0106X
C 6	45.6	11.8	787	29	B2554872
C 7	44.4	11.5	745	28	AY079759
C 8	41.6	10.7	1101	29	CNS000D1
C 9	41.2	10.6	653	28	A2525216
C 10	40.8	10.5	528	28	A2617991
C 11	40.6	10.5	1201	13	BM417132
C 12	40.2	10.4	1101	29	CNS003AL
C 13	39.6	10.2	938	29	CNS006TJ
C 14	39.2	10.1	206	10	BG628372
C 15	39.2	10.1	709	9	AV718352
C 16	39	10.1	512	12	B3334524
C 17	39	10.1	535	12	B3392005
C 18	39	10.1	549	12	B3365679
C 19	39	10.1	604	12	B336012
C 20	39	10.1	638	12	B3396695
C 21	39	10.1	1201	13	BX335036
C 22	39	10.1	1204	29	CNS016E2
C 23	38.8	10.0	396	29	CC455762
C 24	38.8	10.0	1101	29	CNS017KX
C 25	38.8	10.0	1174	9	AL546393
C 26	38.6	10.0	556	12	BM277853
C 27	38.4	9.9	493	12	B3436034
C 28	38.4	9.9	643	9	AU264688
C 29	38.4	9.9	663	28	BH157916
C 30	38.4	9.9	1201	13	BX400522
C 31	38.2	9.9	605	12	BJ414846
C 32	38.2	9.9	664	9	AV865199
C 33	38.2	9.9	677	13	BM283869
C 34	38.2	9.9	722	13	BM154348
C 35	38.2	9.9	761	13	BM201400
C 36	38.2	9.9	770	13	BM068879
C 37	38.2	9.9	1170	13	BX394629
C 38	38	9.8	573	12	B3378220
C 39	38	9.8	635	12	B3373498
C 40	38	9.8	655	12	B3373432
C 41	38	9.8	691	12	B3374478
C 42	38	9.8	718	12	B3373073
C 43	38	9.8	718	12	B3375167
C 44	38	9.8	746	12	B3373989
C 45	38	9.8	748	12	B3375231

ALIGNMENTS

RESULT 1

BH770770/c

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

BH770770 LLMGtag515 MG1363 Random Sequence Tag Library Lactococcus lactis subsp. cremoris genomic, genomic survey sequence.

BH770770 GI:20373727

Lactococcus

1 (bases 1 to 2662)

Boletín A., Ehrlich, S.D. and Sorokin, A.

Studies of genomes of dairy bacteria Lactococcus lactis

Sci. Aliments, (2002) In press

Contact: Sorokin A

Genetique Microbienne

INRA

CRJ INRA, Domaine de Vilvert, 78352 Jouy en Josas cedex, France

Tel: 33 1 34 65 25 16

Fax: 33 1 34 65 25 21

Email: sorokine@jouy.inra.fr

best homologue in strain IL1403 is yphL (98%)

Class: shotgun

High quality sequence start: 30

High quality sequence stop: 2634.

FEATURES
source
Location/Qualifiers
1. 2662
/organism="Lactococcus lactis subsp. cremoris"
/mol_type="genomic DNA"
/strain="MG1363"
/db_xref="taxon:1359"
/clone_lib="MG1363 Random Sequence Tag Library"
/note="Vector: pSGM02; Site_1: SmaI; Library of
chromosomal fragments of L.lactis strain MG1363 was
prepared by partial AluI digestion or by sonication."
BASE COUNT 871 a 506 c 513 g 772 t
ORIGIN

Query Match 41.2%; Score 159.4; DB 28; Length 2662;
Best Local Similarity 63.4%; Pred. No. 4.9e-32;
Matches 244; Conservative 0; Mismatches 141; Indels 0; Gaps 0;

OY 3 TCTTCTTCTTCCACCAAAATGAGAAACAACTGCATCTTAACAAGTCACCAAGACCTAAA 62
DB 2038 TCTTCCGATTTCTATTAGGAAGATTTTGAACAATAGCATCAAGTACATCCCTGTCCCAATC 1979

OY 63 CCATCTGACCTGATATCGGATCGGTTCCACCAAACTCTTAATGAATAGAAATCATACACG 122
DB 1978 CGCTCAACAGCTGATACAGGATATGGATCCCGAGTCCCAAGTGAATAGAAGTCAAAATTT 1919

OY 123 TCTGTACCATTTCCAT 182
DB 1918 TCCATTCGACGTTCGGGGTGTGCGACTTTATGACGATAGAAATGACAGGTTTATCTGTA 1859

OY 183 TTGTATAAAATTTGAGCGACCATTTTCATCGCTTTGTGTCAATCTTTCAGCGACGTTAAC 242
DB 1858 CGATAAAGATATTAGCAACTGCTTCATCGGCATCAGTAATCTCTTCTCCATCAACG 1799

OY 243 ATAAAAATTAATACATCGCTTCATCATCGCATTTCTGCGCTCTTAATTTGTGTT 302
DB 1798 ACAGCGATTAATACATCTGCTCGCTATGCGCATTTGCGCTGCGCGCAATTCAGTC 1739

OY 303 TGGATGTGTCATCACCAGTTCATCAATACCACTGTATCAATATATATATATATATATAT 362
DB 1738 ATGAAAGTTTCATCGAAGTTCATCCACCGGTGTCGATGATGTTGAATTTACGGGTT 1679

OY 363 AACCATTCACCTGAAGAATAAATAC 387
DB 1678 AACCATTCCTGTGGCATAGATAC 1654

RESULT 2
AI390536/c 552 bp mRNA linear EST 15-MAR-2000
LOCUS mu83b02.y1 Stratagene mouse melanoma (#937312) Mus musculus cDNA
DEFINITION clone IMAGE:652107 5' similar to SW:YHPC-BACSU P50743 HYPOTHETICAL
48.8 KD GTP-BINDING PROTEIN IN CMK-GPSA INTERGENIC REGION. ; mRNA
sequence.
ACCESSION AI390536
VERSION AI390536.1 GI:4216543
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 552)
AUTHORS Marra, M., Hillier, L., Kucaba, T., Martin, J., Beck, C., Wylie, T.,
Underwood, K., Steptoe, M., Theising, B., Allen, M., Bowers, Y., Person
B., Swaller, T., Gibbons, M., Pape, D., Harvey, N., Schuck, R., Ritter
E., Kohn, S., Shih, T., Jackson, Y., Cardenas, M., McCann, R.,
Waterson, R., and Wilson, R.
The WashU-NCI Mouse EST Project 1999
Unpublished
TITLE Contact: Marra M/WashU-NCI Mouse EST Project 1999
JOURNAL Washington University School of Medicine
COMMENT 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810

Email: mouseest@watson.wustl.edu
This clone is available royalty-free through LML; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
This read is a RESEQUENCE of a previously sequenced mouse clone
This read has been verified (found to hit its original self in the
correct orientation)
Possible reversed clone: similarity on wrong strand
MG1:397955
Seq primer: -40RP from Gibco
High quality sequence stop: 420
POLYA-No.

FEATURES
source
Location/Qualifiers
1. 552
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone_image:652107
/tissue_type="melanoma"
/dev_stage="M2 cells"
/lab_host="SOLR (kanamycin resistant)"
/clone_lib="Stratagene mouse melanoma (#937312)"
; Site_1: XhoI; Cloned unidirectionally. Primer: Oligo
dT. From M2 cells, a highly metastatic derivative of the
K-1735 (mouse) melanoma. Average insert size: 1.0 kb;
Uni-ZAP XR Vector; -5' adaptor sequence: 5' GAATTCGGCAGCAG
3' -3' adaptor sequence: 5' CTCGAGTTTTTTTTTTTTTTT 3'

BASE COUNT 220 a 63 c 94 g 174 t 1 others

Query Match 24.1%; Score 93.4; DB 9; Length 552;
Best Local Similarity 56.0%; Pred. No. 2e-14;
Matches 200; Conservative 0; Mismatches 151; Indels 6; Gaps 1;

OY 17 ACCAAATGAGAAACAACTGCATCTTAACAAGTCACCAAGACCTTAACACCTGACCTGA 76
DB 351 ATCTAACTCTAATGTTTTCATCTATAATAATCAACACTCTCTTCATGAATAGCTGA 292

OY 77 TATCGGATACGGTTCACCAAAATCTAATGAATAGAAATCATACACTCTGTACGGATTTC 136
DB 291 AATTGGAAAAATATGTTCCAAAACCTAGTGAATAAATTTGAAGATATC-----AAAAAATTT 238

OY 137 CATATTATCTACTTTGTTAAACCGCTAATACGACCGGTTTTTTTAGATTGTTATATAAATTTG 196
DB 237 GTTTCCTCTCTAGTTTATAGCTGGGATTAACAACTTTTTTGTGTTGTTTTCGAAGCAATC 178

OY 197 ACGGACCATTTTCATCGCTTTGTCATCTTCCAGCGACCTTAACCATATAAATAATAAC 256
DB 177 TATACAAATAGTCTCTGATCTAATTTCACTCTACCATCTACTATAAATAATAATTA 118

OY 257 ATCCGCTTCATCTATGGGGATTTCTGCGCTCGGCTCTAATTTGTTTGAATGGTGCATC 316
DB 117 ATTGCTCTCTCTATGCTATTGAGTTTGAATTTGTTTCTTGAATGATTGTTGTT 58

OY 317 ACCAATTTCAATACCACTCTGATCAATAATATTTGAATCATGTTGTTAAACCATTCACC 373
DB 57 TTCTATTCAATTCGCGCTGTCTATCAGATAAATAATTCCTTTCCTGCTCATTCAC 1

RESULT 3
AA215138/c 431 bp mRNA linear EST 03-FEB-1997
LOCUS mu83b02.r1 stratagene mouse melanoma (#937312) Mus musculus cDNA
DEFINITION clone IMAGE:652107 5' similar to TR:G1046032 G1046032 HYPOTHETICAL
50.2 KD PROTEIN. ; mRNA sequence.
ACCESSION AA215138
VERSION AA215138.1 GI:1814868
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 431)

QY 374 TGAAGAATAAA 384
 Db 98 TTCTCCATAAA 88

RESULT 5

CNS0106X/c

LOCUS Drosophila melanogaster genome survey sequence T7 end of BAC
 DEFINITION BACN03K20 of DrosBAC library from Drosophila melanogaster (fruit fly), genomic survey sequence.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

Drosophila melanogaster (fruit fly)

Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;

Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;

Ephydroidea; Drosophilidae; Drosophila.

1 (bases 1 to 1101)

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Direct Submission
 Submitted (23-JUL-1999) Genoscope - Centre National de Sequencage :
 BP 191 91006 EVRY cedex - FRANCE (E-mail : seqrefgenoscope.cns.fr
 - Web : www.genoscope.cns.fr)
 Determination of this BAC-end sequence was carried out as part of a
 collaboration with the European Drosophila Genome Project (EDGP) -
 http://www.edgp.ebi.ac.uk/. This Drosophila melanogaster BAC
 library (Dros BAC) was made by Alain Billaud at CEPH (Centre
 d'Etude du Polymorphisme Humain) with funding provided by a MRC
 project grant. The DNA was prepared from embryos by Alain Bucheton
 and Genevieve Payan. It has been constructed in the vector
 pBelobAC11.

FEATURES

source

1..1101

/organism="Drosophila melanogaster"

/mol_type="genomic DNA"

/db_xref="taxon:7227"

/clone="BACN03K20"

/plasmid="pBelobAC11"

/note="end : T7"

BASE COUNT 258 a 107 c 60 g 175 t 501 others

ORIGIN

Query Match

Best Local Similarity 12.0% Score 46.6; DB 29; Length 1101;

Matches 71; Conservative 155; Mismatches 159; Indels 0; Gaps 0;

QY 2 ATCTTCTCTCTTCCACCAAAATGAGAAACAATCTCATCTAACCAAGTCACCAAGACCTAA 61

Db 974 MTHTHHH 915

QY 62 ACCATGTGACCTTGATACGGATCGGTTCACCAATCTTAATGAAGAATCATACAC 121

Db 914 TMTTHHH 855

QY 122 GTCGTACCATTTCCATATATCTACTTGTTCACCCCTAATACGACCGGTTTTTGA 181

Db 854 YMTMTTCTTTTMTTTHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHH 795

QY 182 TTGTATATAAATTTGAGCGACCATTTTCATCGCTTTTGTCTCAATCTTCACGCGTTAAC 241

Db 794 TMTHTTHTHTHTHTHTTMTTMTTMTTMTTMTTMTTMTTMTTMTTMTTMTTMTT 735

QY 242 CATATAAATAAATACATCGCTTCATCTATGCGGATTTCTCGCTCGCTCTAATTTGT 301

Db 734 HTTATATATATATTTTMTTMTTMTTMTTMTTMTTMTTMTTMTTMTTMTTMTT 675

QY 302 TTGGAATGGTCATACCAATTTCAATACCACTCTATCAATATATTTGAATCATGT 361

Db 674 WMTHTTHTTHTTHTTMTTMTTMTTMTTMTTMTTMTTMTTMTTMTTMTTMTT 615

QY 362 TAACCATTACCTGAAGAATAAATA 386
 Db 614 MTHHAACACACMTWAKAAAWAA 590

RESULT 6

B2554672/c

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

Pseudomonas aeruginosa

Pseudomonas aeruginosa

Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;

1 (bases 1 to 787)

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Whole-Genome-Sequence variation among multiple isolates of

Pseudomonas aeruginosa library

J. Bacteriol. (2002) In press

Contact: Chris K. Raymond

Genome Center

University of Washington

Box 352145, Seattle, WA 98105-2145, USA

Tel: 2062216954

Fax: 2066857244

Email: craymond@u.washington.edu

Class: shotgun.

FEATURES

source

1..787

/organism="Pseudomonas aeruginosa"

/mol_type="genomic DNA"

/strain="1-60"

/db_xref="taxon:287"

/clone="pacsl-60_4804"

/note="clinical isolate 1-60 Whole genomic shotgun library."

BASE COUNT 180 a 229 c 251 g 123 t 4 others

ORIGIN

Query Match

Best Local Similarity 11.8% Score 45.6; DB 29; Length 787;

Matches 129; Conservative 0; Mismatches 139; Indels 0; Gaps 0;

QY 78 ATCGGATACGGTTCCACCAATCTTAATGAATAGAAATACACGCTGTGACGATTTCC 137

Db 515 ATGGCGACGCTTGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 456

QY 138 ATATTATCTACTTTTGTAAACCGCTAATACGACGGTTTTTTTATAGATTTTGTATAAATTTGA 197

Db 455 ATGTTGTGCGCTTGTGCGGATCAGGAAGTTGCCCTTGTCCGCTGCCAGGTGCTCG 396

QY 198 CGGACATTTTCATCGCTTTGTGTCATCTCTTCACGACGCTTAACCATATAAATAAATA 257

Db 395 CCGATCATCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 336

QY 258 TCCGCTTCATCTATGCGGATTTCTGCTCGCTCTAATTTGTGTTGGATGTTGTCATCA 317

Db 335 TCCGCTTCCTCGATGGCTTGCACGATTCCTCGGCGGCGGCGGCGGCGGCGGCGGCGG 276

QY 318 CCAATTTCAATACCATCTGATCAATAA 345

Db 275 TCGCGGAGATGCTCGGATTCGATGA 248

RESULT 7

AY079759/c

LOCUS

DEFINITION

AY079759

Scripps Pier

745 bp

DNA

linear

GSS 06-NOV-2002

AY079759

AY079759

Scripps Pier (La Jolla, CA) uncultured virus community


```

BX417132
LOCUS      BX417132      1201 bp      mRNA      linear      EST 13-MAY-2003
DEFINITION BX417132 Homo sapiens PLACENTA Homo sapiens cDNA clone CS0DE004YH12
ACCESSION  BX417132
VERSION     BX417132.1 GI:30633228
KEYWORDS   EST.
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens
REFERENCE  1 (bases 1 to 1201)
AUTHORS   Li M.B., Gruber,C., Jesse,J. and Pollayes,D.
JOURNAL   Full-length cDNA libraries and normalization
COMMENT   Contact: Feng Liang Email: fliang@lifetech.com URL: 2242.r
          http://fulllength.invitrogen.com/ invitrogen Corporation 1600
          Faraday Avenue Genoscope sequence ID: CS0DE004DD06NP1.
          Location/Qualifiers
FEATURES   source
            1..1201
            /organism="Homo sapiens"
            /mol_type="mRNA"
            /db_xref="taxon:9606"
            /clone="CS0DE004YH12"
            /tissue_type="PLACENTA"
            /clone_lib="Homo sapiens PLACENTA"
            /note="Vector: pCMVSPORT1.6; 1st strand cDNA was primed
            with a NotI-oligo(dT) primer. Five prime end enriched,
            double-strand cDNA was digested with Not I and cloned into
            the Not I and EcoRV sites of the pCMVSPORT 6 vector.
            Library was not normalized."
BASE COUNT 359 a 221 c 212 g 245 t 164 others
ORIGIN
Query Match 10.5%; Score 40.6; DB 13; Length 1201;
Best Local Similarity 31.1%; Pred. No. 4;
Matches 75; Conservative 55; Mismatches 111; Indels 0; Gaps 0:
QY 17 ACCAAATGAGAACAACTGCATCTAACAGTGCACCAAGACCTAACACCATGTGACCTGA 76
   |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:|
Db 826 AMGAAAAAANAANAAGNNKWNMAAANAANAANAANAANAANAANAANAANAANAANA
   |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:|
QY 77 TATCGGATACGGTTCACCAATCCTAATGAATAGAAATCATACACGCTGTGACGATTTC 136
   |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:|
Db 886 GGGKGGGCKKRRTAAGCAANTYATAWAANAANAANAANAANAATTAACCAANAANAATM 945
   |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:|
QY 137 CATATTATCTACTTTGTTAACCGCTAATACGACCGGTTTTTTAGATTGTTATAAATTTG 196
   |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:|
Db 946 HAAATTTTATTTAAHAHAATWTTAAANAATVTTATTTTTTAAANAANAANAANAATWAAT 1005
   |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:|
QY 197 AGCGACCATTTTCGCTTGTGTCATCCTTCACGACGCTTAACCATATAAATAAATAAC 256
   |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:|
Db 1006 TTTTAAATTAAYTATAAYVYANCHYVYMAAAYACVMAAAYCAAMWATTAAYC 1065
   |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:|
QY 257 A 257
Db 1066 A 1066

RESULT 12
CNS003AL
LOCUS      CNS003AL      1101 bp      DNA      linear      GSS 03-JUN-1999
DEFINITION Drosophila melanogaster genome survey sequence T7 end of BAC #
            BACR08K13 of RPCI-98 library from Drosophila melanogaster (fruit
            fly), genomic survey sequence.
ACCESSION  AL063962
VERSION     AL063962.1 GI:4941819
KEYWORDS   GSS.
SOURCE     Drosophila melanogaster (fruit fly)

BX417132
LOCUS      BX417132      1201 bp      mRNA      linear      EST 13-MAY-2003
DEFINITION BX417132 Homo sapiens PLACENTA Homo sapiens cDNA clone CS0DE004YH12
ACCESSION  BX417132
VERSION     BX417132.1 GI:30633228
KEYWORDS   EST.
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens
REFERENCE  1 (bases 1 to 1201)
AUTHORS   Li M.B., Gruber,C., Jesse,J. and Pollayes,D.
JOURNAL   Full-length cDNA libraries and normalization
COMMENT   Contact: Feng Liang Email: fliang@lifetech.com URL: 2242.r
          http://fulllength.invitrogen.com/ invitrogen Corporation 1600
          Faraday Avenue Genoscope sequence ID: CS0DE004DD06NP1.
          Location/Qualifiers
FEATURES   source
            1..1201
            /organism="Homo sapiens"
            /mol_type="mRNA"
            /db_xref="taxon:9606"
            /clone="CS0DE004YH12"
            /tissue_type="PLACENTA"
            /clone_lib="Homo sapiens PLACENTA"
            /note="Vector: pCMVSPORT1.6; 1st strand cDNA was primed
            with a NotI-oligo(dT) primer. Five prime end enriched,
            double-strand cDNA was digested with Not I and cloned into
            the Not I and EcoRV sites of the pCMVSPORT 6 vector.
            Library was not normalized."
BASE COUNT 359 a 221 c 212 g 245 t 164 others
ORIGIN
Query Match 10.5%; Score 40.6; DB 13; Length 1201;
Best Local Similarity 31.1%; Pred. No. 4;
Matches 75; Conservative 55; Mismatches 111; Indels 0; Gaps 0:
QY 17 ACCAAATGAGAACAACTGCATCTAACAGTGCACCAAGACCTAACACCATGTGACCTGA 76
   |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:|
Db 826 AMGAAAAAANAANAAGNNKWNMAAANAANAANAANAANAANAANAANAANAANAANA
   |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:|
QY 77 TATCGGATACGGTTCACCAATCCTAATGAATAGAAATCATACACGCTGTGACGATTTC 136
   |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:|
Db 886 GGGKGGGCKKRRTAAGCAANTYATAWAANAANAANAANAANAATTAACCAANAANAATM 945
   |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:|
QY 137 CATATTATCTACTTTGTTAACCGCTAATACGACCGGTTTTTTAGATTGTTATAAATTTG 196
   |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:|
Db 946 HAAATTTTATTTAAHAHAATWTTAAANAATVTTATTTTTTAAANAANAANAANAATWAAT 1005
   |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:|
QY 197 AGCGACCATTTTCGCTTGTGTCATCCTTCACGACGCTTAACCATATAAATAAATAAC 256
   |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:|
Db 1006 TTTTAAATTAAYTATAAYVYANCHYVYMAAAYACVMAAAYCAAMWATTAAYC 1065
   |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:|
QY 257 A 257
Db 1066 A 1066

RESULT 12
CNS003AL
LOCUS      CNS003AL      1101 bp      DNA      linear      GSS 03-JUN-1999
DEFINITION Drosophila melanogaster genome survey sequence T7 end of BAC #
            BACR08K13 of RPCI-98 library from Drosophila melanogaster (fruit
            fly), genomic survey sequence.
ACCESSION  AL063962
VERSION     AL063962.1 GI:4941819
KEYWORDS   GSS.
SOURCE     Drosophila melanogaster (fruit fly)

```

```

KEYWORDS   GSS.
SOURCE     Drosophila melanogaster (fruit fly)
ORGANISM   Drosophila melanogaster
REFERENCE  1 (bases 1 to 1101)
AUTHORS   Aaron Mammor in Pieter de Jong's laboratory in the Department of
            Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,
            NY. The library is named RPCI-98 and was constructed by partial
            EcoRI digestion of Drosophila DNA provided by the BDGP from the
            isogenic strain Y2; cn bw sp, the same strain used for the BDGP's
            P1 and EST libraries. A more detailed description of the library
            and how to order individual BAC clones, the entire library, or
            filters for hybridization from the BACPAC Resource Center can be
            found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.
            Location/Qualifiers
FEATURES   source
            1..1101
            /organism="Drosophila melanogaster"
            /mol_type="genomic DNA"
            /db_xref="taxon:7227"
            /clone="BACR08K13"
            /clone_lib="RPCI-98"
            /note="end: T7"
BASE COUNT 303 a 195 c 173 g 280 t 150 others
ORIGIN
Query Match 10.4%; Score 40.2; DB 29; Length 1101;
Best Local Similarity 28.3%; Pred. No. 5;
Matches 80; Conservative 72; Mismatches 131; Indels 0; Gaps 0:
QY 20 AAAATGAGAACAACTGCATCTAACAGTGCACCAAGACCTAACACCATGTGACCTGATAT 79
   |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:|
Db 819 AAAAACHAAHAHAANAANAANAANAANAANAANAANAANAANAANAANAANAANA
   |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:|
QY 80 CGGATACGGTTCACCAAAATCCTAATGAATAGAAATCATACACGCTGTGACCATTTCCAT 139
   |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:|
Db 879 TYMHHTTTTATTTHAANAANAANAANAANAANAANAANAANAANAANAANAANA
   |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:|
QY 140 ATTATCTACTTTGTTAACCGCTAATACGACCGGTTTTTTAGATTGTTATAAATTTGAGC 199
   |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:|
Db 939 ACYHHHMHNAACHAYCTTCYCTWAWHCAATWTTTTHWTTTTHWTTTTHWTTTTHWTTT 998
   |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:|
QY 200 GACCATTTTCATCGCTTTGTGTCATCCTTCACGACGCTTAACCATATAAATAAATACATC 259
   |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:|
Db 999 AWTATTTTTTTTTTTTTTTTAAATATATCTHNAWMTCTHTWYVATCTNMVACACACAYA 1058
   |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:|
QY 260 CGCTTCATCTATCGCGATTTCTCGCTCGCTCTAAATTTGTGTT 302
   |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:|
Db 1059 TMWYHYVTHHTTATTTTYYHTCATAYVATTATATCTCMWT 1101
   |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:|

RESULT 13
CNS006TJ
LOCUS      CNS006TJ      938 bp      DNA      linear      GSS 03-JUN-1999
DEFINITION Drosophila melanogaster genome survey sequence T7 end of BAC #
            BACR14J11 of RPCI-98 library from Drosophila melanogaster (fruit
            fly), genomic survey sequence.
ACCESSION  AL065906
VERSION     AL065906.1 GI:4944874
KEYWORDS   GSS.
SOURCE     Drosophila melanogaster (fruit fly)

```


